

1 QUINN EMANUEL URQUHART & SULLIVAN, LLP
2 Alex Spiro (admitted *pro hac vice*)
3 alexspiro@quinnemanuel.com
4 51 Madison Avenue, 22nd Floor
New York, New York 10010
Telephone: (212) 849-7000

5 QUINN EMANUEL URQUHART & SULLIVAN, LLP
6 Robert M. Schwartz (Bar No. 117166)
7 robertschwartz@quinnemanuel.com
8 Michael T. Lifrak (Bar No. 210846)
9 michaellifrak@quinnemanuel.com
10 Jeanine M. Zalduendo (Bar No. 243374)
11 jeaninezalduendo@quinnemanuel.com
12 865 South Figueroa Street, 10th Floor
13 Los Angeles, California 90017-2543
14 Telephone: (213) 443-3000

15 *Attorneys for Defendant Elon Musk*

16
17
18
19
20
21
22
23
24
25
26
27
28

**UNITED STATES DISTRICT COURT
CENTRAL DISTRICT OF CALIFORNIA**

VERNON UNSWORTH,

Plaintiff,

vs.

ELON MUSK,

Defendant.

Case No. 2:18-cv-08048

Judge: Hon. Stephen V. Wilson

**SUPPLEMENTAL DECLARATION
OF MICHAEL T. LIFRAK IN
SUPPORT OF DEFENDANT'S
MOTION IN LIMINE NO 5 TO
EXCLUDE THE EXPERT OPINION
OF DR. BERNARD J. JANSEN**

Complaint Filed: September 17, 2018
Trial Date: December 2, 2019

Hearing Date: November 25, 2019
Time: 3:00 p.m.
Courtroom: 10A

1 **I, Michael T. Lifrak, declare as follows:**

2 1. I am a member of the bar of the State of California and a partner at
3 Quinn Emanuel Urquhart & Sullivan, LLP, attorneys for Defendant Elon Musk. I
4 make this declaration of personal, firsthand knowledge, and if called and sworn as a
5 witness, I could and would testify competently thereto.

6 2. I submit this supplemental declaration in support of Mr. Musk's Motion
7 in Limine No. 5 to Exclude the Expert Opinion of Dr. Bernard J. Jansen.

8 3. Attached hereto as **Exhibit 8** is a true and correct copy of the front page
9 of the Pictorial City Sheet from the February 6, 1915 edition of the *Los Angeles*
10 *Times*, available at www.newspapers.com/image/380140241 (last visited Nov. 20,
11 2019).

12 4. Attached hereto as **Exhibit 9** is a true and correct copy of the cover of
13 the September 23, 1969 issue of *Look Magazine*, available at
14 www.oldlifemagazines.com/look-magazine-september-23-1969-diana-ross.html (last
15 visited Nov. 20, 2019.).

16 5. Attached hereto as **Exhibit 10** is a true and correct copy of *New York*
17 *Times Article Archive*, available at
18 <https://archive.nytimes.com/www.nytimes.com/ref/membercenter/nytarchive.html>
19 (last visited Nov. 21, 2019).

20 6. Attached hereto as **Exhibit 11** is a true and correct copy of the *July 15,*
21 *2018 Home Page of Foxnews.com*, WAYBACK MACHINE-ARCHIVE,
22 <https://web.archive.org/web/20180715205857/http://www.foxnews.com/> (last visited
23 Nov. 20, 2019).

24 7. Attached hereto as **Exhibit 12** is a true and correct copy of Heilman, J.
25 (2015) "Open Access to a High-Quality, Impartial, Point-of-Care Medical Summary
26 Would Save Lives: Why Does It Not Exist?" PLoS Med 12(8): e1001868, available
27 at <https://doi.org/10.1371/journal.pmed.1001868> (last visited Nov. 21, 2019).
28

1 8. Attached hereto as **Exhibit 13** is a true and correct copy of Szolnoki, P.,
2 et al, “Who Should be my Facebook Partner? Analysis of the Relationship between
3 Hungarian Large-scale Facebook Pages,” *Procedia Computer Science*, Vol. 101:86-
4 95 (2016), available at
5 <https://www.sciencedirect.com/science/article/pii/S1877050916326795> (last visited
6 Nov. 19, 2019).

7 9. Attached hereto as **Exhibit 14** is a true and correct copy of Baumel, A.,
8 PhD, et al., “Objective User Engagement with Mental Health Apps: Systematic
9 Search and Panel-Based Usage Analysis,” *J. Med. Internet Res.* 2019;21(9):e14567,
10 available at <https://www.jmir.org/2019/9/e14567/> (last visited Nov. 21, 2019).

11 10. Attached hereto as **Exhibit 15** is a true and correct copy of
12 Brumshteyn, Y., “Analysis of the Webometric Indicators of the Main Websites that
13 Aggregate Multithematic Scientific Information,” *Automatic Documentation &*
14 *Mathematical Linguistics* (Nov. 2017) Vol. 51:6:250-265 available at
15 <https://doi.org/10.3103/S0005105517060048> (last visited Nov. 21, 2019).

16 11. Attached hereto as **Exhibit 16** is a true and correct copy of *Similar Web*
17 *vs. Google Analytics*, which was copied from [https://support.similarweb.com/hc/en-](https://support.similarweb.com/hc/en-us/articles/360001638917-SimilarWeb-vs-Google-Analytics)
18 [us/articles/360001638917-SimilarWeb-vs-Google-Analytics](https://support.similarweb.com/hc/en-us/articles/360001638917-SimilarWeb-vs-Google-Analytics) on November 2, 2019.

19 12. Attached hereto as **Exhibit 17** is a true and correct copy of *SimilarWeb*
20 *vs. Direct Measurement*, available at [https://www.similarweb.com/blog/wp-](https://www.similarweb.com/blog/wp-content/uploads/2016/08/SW-vs-Direct-Measurement.pdf)
21 [content/uploads/2016/08/SW-vs-Direct-Measurement.pdf](https://www.similarweb.com/blog/wp-content/uploads/2016/08/SW-vs-Direct-Measurement.pdf) (last visited Nov. 21,
22 2019).

23 13. Attached hereto as **Exhibit 18** is a true and correct copy of the Excel
24 spreadsheet titled “Links for Classification” which was produced as part of a larger
25 Excel Workbook by Dr. Jansen at JANSEN08202.

26 14. Attached hereto as **Exhibit 19** is a true and correct copy of the front
27 page of the Telegraph Sheet from the February 6, 1915 edition of the *Los Angeles*
28

1 *Times*, available at <https://www.newspapers.com/image/380140106> (last visited Nov.
2 21, 2019).

3 15. Attached hereto as **Exhibit 20** is a true and correct copy of Exhibit 145
4 to the November 4, 2019 deposition of Dr. Bernard “Jim” Jansen, SimilarWeb
5 Website Performance Report for foxnews.com, produced at JANSEN_05554-58.

6 16. Attached hereto as **Exhibit 21** is a true and correct copy of the
7 SimilarWeb Website Performance Report for abcnews.go.com, produced at
8 JANSEN_06407-11.

9 17. Attached hereto as **Exhibit 22** is a true and correct copy of the
10 SimilarWeb Website Performance Report for cbsnews.com, produced at
11 JANSEN_07177-81.

12
13 I declare under penalty of perjury under the laws of the State of California that
14 the foregoing is true and correct and that this document was executed in Los
15 Angeles, California.

16
17 DATED: November 21, 2019

18
19 By 

20 Michael T. Lifrak
21
22
23
24
25
26
27
28

EXHIBIT 8

Exhibit 8, Page 4

EXHIBIT 9

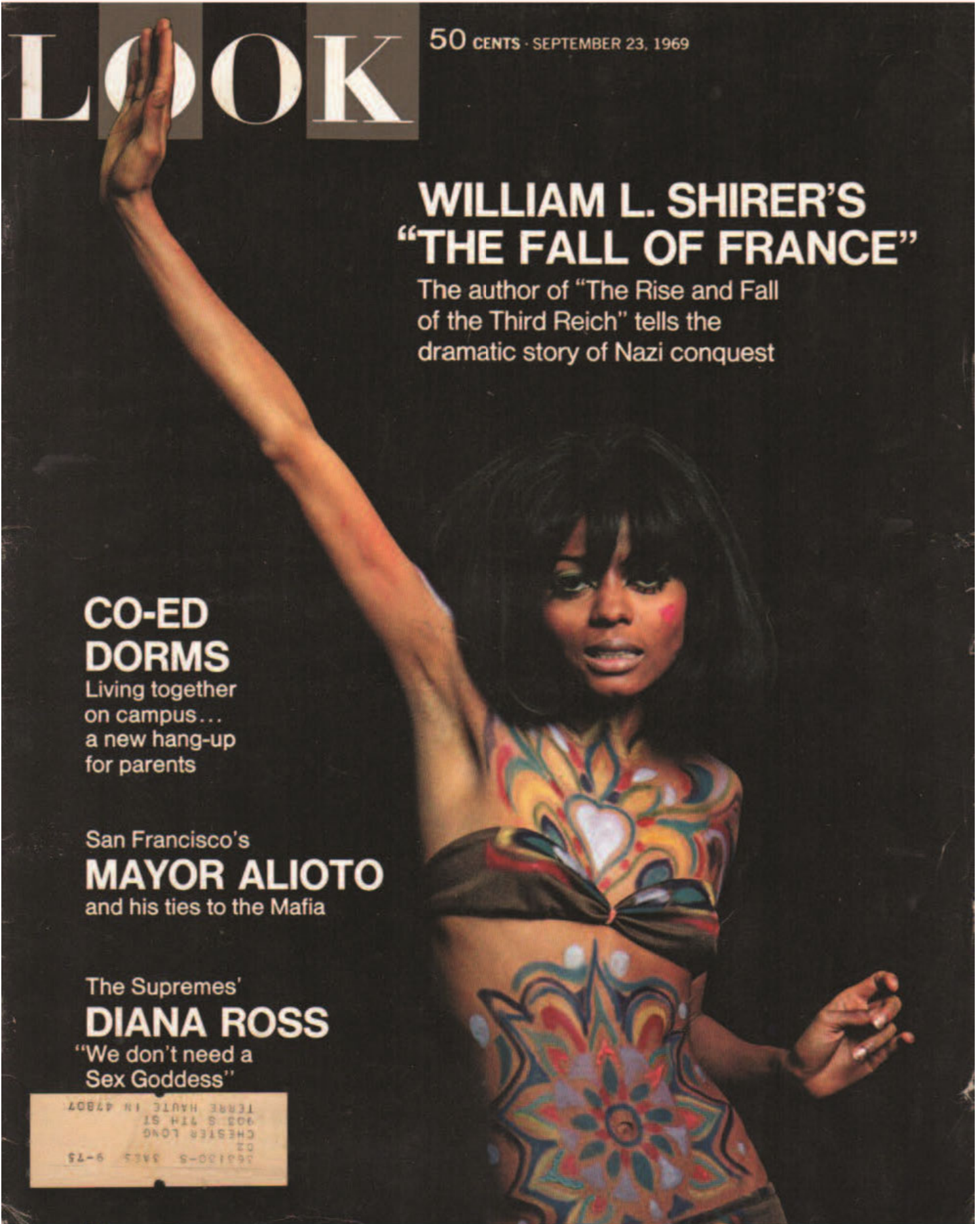


EXHIBIT 10

This is an archived page. [Report a problem](#)

[HOME PAGE](#) [TODAY'S PAPER](#) [VIDEO](#) [MOST POPULAR](#) [U.S. Edition](#)

The New York Times

Archive

Search All NYTimes.com

Go

[WORLD](#) [U.S.](#) [N.Y. / REGION](#) [BUSINESS](#) [TECHNOLOGY](#) [SCIENCE](#) [HEALTH](#) [SPORTS](#) [OPINION](#) [ARTS](#) [STYLE](#) [TRAVEL](#) [JOBS](#) [REAL ESTATE](#) [AUTOS](#)

New York Times Article Archive

1851–PRESENT

The complete archive of The New York Times can now be searched from NYTimes.com — more than 13 million articles total.

Searching the Archives

The archive is divided into two search sets: 1851–1980 and 1981–present.

[Search the Article Archive: 1981–Present »](#)

[Search the Article Archive: 1851–1980 »](#)

Accessing and Purchasing Articles

Digital Subscribers:

— **1923–1980:** Your digital subscription includes 100 archive articles every four weeks in this date range (from January 1, 1923 through December 31, 1980). After you've reached the 100-article limit for the month, articles from 1923 through 1980 are \$3.95 each.

— **Pre-1923 and post-1980:** Articles published before January 1, 1923 or after December 31, 1980 are free with your digital subscription and are not limited in any way.

[Learn more about digital subscriptions »](#)

Nonsubscribers:

— **1923–1980:** Articles in this date range (from January 1, 1923 through December 31, 1980) are available for purchase at \$3.95 each.

— **Pre-1923 and post-1980:** Articles published before January 1, 1923 or after December 31, 1980 are free, but they count toward your monthly limit.

FACEBOOK

TWITTER

GOOGLE+

SAVE

EMAIL

SHARE

PRINT

REPRINTS

[Learn more about your monthly limit as a nonsubscriber »](#) #4920

Users of NYTimes.com Passes:

- 1923–1980: Pass users have unlimited access to the New York Times archived articles outside the 1923–1980 date range. Each day of their pass, users may access up to five free articles published between the years 1923 through 1980.
- Pre-1923 and post-1980: Articles published before January 1, 1923 or after December 31, 1980 are free with your Pass and are not limited in any way

[Learn more about NYTimes.com Passes »](#)

Members of a Group Subscription:

- 1923–1980: Each member of a group digital subscription can access 100 archive articles every four weeks in this date range (from January 1, 1923 through December 31, 1980).
- Pre-1923 and post-1980: Articles published before January 1, 1923 or after December 31, 1980 are free to members of a group digital subscription and are not limited in any way.

[Learn more about group subscriptions »](#)

Formats

Most articles are available as text only. Photos are available for purchase by e-mailing our photo sales department at photosales@nytimes.com.

Articles from 1851–1922 are also available as [images of pages from the newspaper](#). These images are available to digital subscribers only.

More Information

Archive purchases are non-refundable.

For full details, see the [Archives FAQ](#).

SAVE EMAIL SHARE


This is an archived page. [Report a problem](#)

EXHIBIT 11

<http://www.foxnews.com/>

98,310 captures
22 Oct 1996 – 21 Nov 2019

JUN JUL AUG
◀ 15 ▶
2017 2018 2019



About this capture

Fox News

- [U.S.](#)
- [World](#)
- [Opinion](#)
- [Politics](#)
- [Entertainment](#)
- [Business](#)
- [Lifestyle](#)
- [TV](#)
- [Radio](#)
- [More](#)

[Expand / Collapse search](#)

[Login](#)

[Watch TV](#)



Hot Topics

- [Trump slams EU](#)
- [Brutal cop killing](#)
- [Tsunami threat](#)

Markets

- [SP500](#)
- [I:COMP](#)
- [I:DJ](#)
- [More](#)

Sponsored by

[U.S.](#)

- [Crime](#)
- [Military](#)
- [Education](#)
- [Terror](#)
- [Immigration](#)
- [Economy](#)
- [Personal Freedoms](#)
- [Fox News Investigates](#)

[World](#)

- [U.N.](#)
- [Conflicts](#)
- [Terrorism](#)
- [Disasters](#)
- [Global Economy](#)
- [Environment](#)
- [Religion](#)
- [Scandals](#)

[Politics](#)

- [Executive](#)
- [Senate](#)
- [House](#)
- [Judiciary](#)
- [Foreign policy](#)
- [Polls](#)
- [Elections](#)

[Entertainment](#)

- [Celebrity News](#)
- [Movies](#)
- [TV News](#)
- [Music News](#)
- [Style News](#)
- [Entertainment Video](#)

[Business](#)

- [Markets](#)
- [Politics](#)
- [Technology](#)
- [Features](#)

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98.310 captures

22 Oct 1996 - 21 Nov 2019

- [Food + Drink](#)
- [Cars + Trucks](#)
- [Travel + Outdoors](#)
- [House + Home](#)
- [Fitness + Well-being](#)
- [Style + Beauty](#)

[Science](#)

- [Archaeology](#)
- [Air & Space](#)
- [Planet Earth](#)
- [Wild Nature](#)
- [Natural Science](#)
- [Dinosaurs](#)

[Tech](#)

- [Security](#)
- [Innovation](#)
- [Drones](#)
- [Computers](#)
- [Video Games](#)
- [Military Tech](#)

[Health](#)

- [Healthy Living](#)
- [Medical Research](#)
- [Mental Health](#)
- [Cancer](#)
- [Heart Health](#)
- [Children's Health](#)

[TV](#)

- [Shows](#)
- [Personalities](#)
- [Watch Live](#)
- [Full Episodes](#)
- [Show Clips](#)
- [News Clips](#)

[About](#)

- [Careers](#)
- [College Students](#)
- [Fox Around the World](#)
- [Advertise With Us](#)
- [Ad Choices](#)
- [Email Newsroom](#)
- [Media Relations](#)
- [Compliance](#)

[Other](#)

- [Fox News Insider](#)
- [Fox News Radio](#)
- [Fox Nation](#)
- [Fox News Go](#)
- [Newsletters](#)
- [Alerts](#)
- [Podcasts](#)
- [Apps & Products](#)

[Fox News](#)

- [Terms of Use](#)
- [Privacy Policy](#)
- [Closed Captioning Policy](#)
- [Help](#)

This material may not be published, broadcast, rewritten, or redistributed. ©2018 FOX News Network, LLC. All rights reserved. All market data delayed 20 minutes.

- [Facebook](#)
- [Twitter](#)
- [Google+](#)
- [Instagram](#)
- [RSS](#)
- [Email](#)

<http://www.foxnews.com/>

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98.310 captures

22 Oct 1996 - 21 Nov 2019

channel

PreviewOn

Switch preview on/off

[On Now](#)

Exclusive Video



11:21

[fox news sunday](#)

[Amb. Jon Huntsman previews the Trump-Putin summit](#)



6:09

[justice with judge jeanine](#)

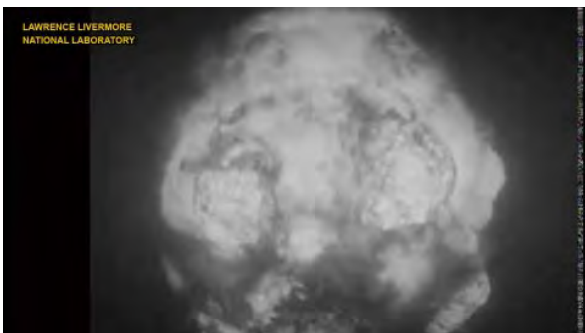
[Rep. Darrell Issa talks explosive Peter Strzok hearing](#)



9:44

[Watters' World](#)

[President Trump brings 'America First' message to Europe](#)



2:35

[Never-before-seen Cold War videos declassified](#)

<http://www.foxnews.com/>

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98,310 captures

22 Oct 1996 - 21 Nov 2019



2:56

[Fox & Friends](#)

'Fox & Friends' celebrates National Ice Cream Day



1:21

Georgia animal control officer rescues husky from hot car



3:40

A look back at Hollywood stars in swimsuits



0:52

Paul Ryan's SUV allegedly eaten by woodchucks

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98,310 captures

22 Oct 1996 - 21 Nov 2019



4:05

Teens with autism become digital media producers



4:02

cavuto live

Former AG Mukasey on Mueller's Russia probe

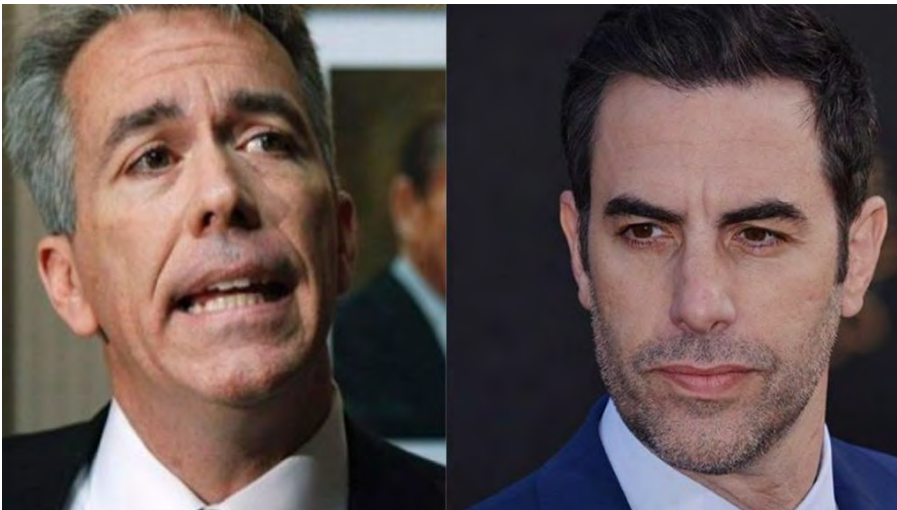
Show More



TOP-LEVEL TENSION

Five sticking points that could spell trouble for Trump, Putin in major summit

- [Terror expert compares reaction of Trump's Russia efforts to Obama's: 'It's glaring hypocrisy'](#)
- [Reagan aide on the Trump-Putin Summit: Here is the one crucial thing to focus on before anything else](#)
- [Trump 'highly unlikely' to recognize Putin's annexation of Crimea, Ambassador Huntsman says](#)
- [Trump says 'European Union is a foe' to US ahead of summit with Putin](#)
- [Bolton says Russian hacking indictment gives Trump upper hand in meeting with Putin](#)
- [Trump advised May to sue EU over Brexit, prime minister reveals](#)



BORAT'S BLUFF

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98,310 captures

22 Oct 1996 - 21 Nov 2019

- [Sacha Baron Cohen reportedly dupes Ted Koppel, Roy Moore, Joe Arpaio; responds to Palin criticism](#)



DEADLY ATTACK

[Massachusetts cop hit with rock, shot several times with own gun: officials](#)

- [Officers killed in the line of duty in 2018](#)
- [New York state trooper killed responding to domestic dispute, official says](#)
- [3 Kansas City, Missouri, cops shot; gunman killed in standoff with police, officials say](#)



VACATION DANGER

[More than 500 fall ill after visiting Tennessee zipline attraction](#)

- [What is bacterial meningitis?](#)
- [Infant dies day after showing symptoms of meningitis, family says](#)



FROZEN FEAR

[11-million-ton iceberg threatens to inundate tiny village with tsunami](#)

- [Mexico earthquake uncovers ancient temple hidden in damaged pyramid](#)
- [Bloodstains on Shroud of Turin are probably fake, experts say](#)



PreviewOn
Switch preview on/off
[On Now](#)

Exclusive Video



11:21

[fox news sunday](#)

[Amb. Jon Huntsman previews the Trump-Putin summit](#)



6:09

[justice with judge jeanine](#)

[Rep. Darrell Issa talks explosive Peter Strzok hearing](#)

<http://www.foxnews.com/>

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

[98,310 captures](#)

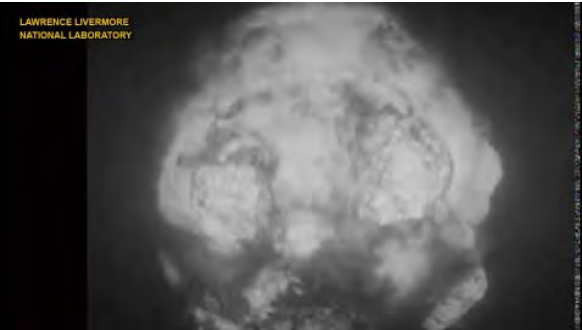
22 Oct 1996 - 21 Nov 2019



9:44

[Watters' World](#)

[President Trump brings 'America First' message to Europe](#)



2:35

[Never-before-seen Cold War videos declassified](#)



2:56

[Fox & Friends](#)

['Fox & Friends' celebrates National Ice Cream Day](#)



1:21

[Georgia animal control officer rescues husky from hot car](#)

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98,310 captures

22 Oct 1996 - 21 Nov 2019



3:40

[A look back at Hollywood stars in swimsuits](#)



0:52

[Paul Ryan's SUV allegedly eaten by woodchucks](#)



4:05

[Teens with autism become digital media producers](#)



4:02

[cavuto live](#)

[Former AG Mukasey on Mueller's Russia probe](#)

[Show More](#)

[U.S.](#)

[3 people killed, 22 injured after multi-vehicle crash involving tour bus in New Mexico](#)

[Politics](#)

[Trumps blasts FBI Agent Strzok as 'disgrace to our country'](#)

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98,310 captures

22 Oct 1996 - 21 Nov 2019

[Trump sees no challenger who could beat him in 2020 as Dems begin to test waters](#)

[Science](#)

[Archaeologists in Egypt discover ancient mummification workshop](#)

[U.S.](#)

[Colorado cops arrest man with head tattoo who was sought in deadly shooting of woman he knew](#)

[Politics](#)

[Defeated Rep. Crowley tries easing tensions with Ocasio-Cortez, won't endorse Pelosi](#)

[U.S.](#)

[Philadelphia-area mom accused of killing her infant through drugs in breast milk](#)

[Entertainment](#)

[Meghan Markle's father Thomas believes his daughter is 'terrified' adjusting to new life, calls royal family 'outdated'](#)

[Politics](#)

[Missouri governor signs law banning marriage of 15-year-olds](#)

[World](#)

[Inside Israel's raid to seize nuclear documents in Iran](#)

[U.S.](#)

['Family shootout' in Maine leads to arrests of 64-year-old man, two sons](#)

[U.S.](#)

[Illegal immigrant who allegedly slashed his wife with a chainsaw was previously deported 11 times: officials](#)

[World](#)

[Anti-Trump paraglider who flew meters away from president arrested](#)

[Features](#)

[Texas man gets prison, must repay \\$2.4M in pay phone scam](#)



Editor's Picks

[Why Princess Diana's niece doesn't mind being in the spotlight](#)

[Entertainment](#)

[Age-progression video released of missing infant snatched 32 years ago](#)

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98,310 captures

22 Oct 1996 - 21 Nov 2019

[World](#)

[Microsoft calls for facial recognition technology rules given 'potential for abuse'](#)

[Tech](#)

[This Honda lawn mower will go 150 mph](#)

[Features](#)

['It's standard': Owens slams Hillary for 'ridiculous' rhetoric on Trump's SCOTUS pick](#)

[Fox News Insider](#)

Opinion

[Trump-Putin summit could improve US-Russia relations](#)

By [Harry J. Kazianis](#)

[Will Hillary persist in 2020? Her actions suggest 'yes'](#)

By [Lauren DeBellis Appell](#)

[Media revel in London anti-Trump protests as obsession with Russia continues](#)

By [Dan Gainor](#)

[Raise a glass to GOP tax cuts! You can literally drink up the benefits](#)

By [Abigail Marone](#)

[Twitter is the latest battlefield between the US and Iran](#)

By [Gregory Keeley](#)

[Go forward in what God is leading you to do in life](#)

By [Ronnie Floyd](#)

[I was Kavanaugh's law clerk. Let me tell you about the man I know](#)

By [Amit Agarwal](#)

[The 12 things Justin must do if he wants his marriage to work](#)

By [Bruce and Lauren Ashford](#)

[Show More](#)

[Podcast](#)

[Fox News Rundown](#)

-



Play / Pause

/

Volume



Get the best of Fox News delivered to your inbox daily

Enter email address

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98,310 captures

22 Oct 1996 - 21 Nov 2019

Close

Features & Faces

[Odd way to cook a steak?](#)

[Are these the BEST beaches?](#)

[Cohen's gun-rights stunt](#)

[Lauer's divorce details](#)

[Lawyer threatens Woods](#)

[Star explains why she disappeared from Hollywood](#)

[Pioneer Woman always keeps this frozen](#)

[Ryan: Woodchucks ate my truck](#)

[What's the best streaming service?](#)

[Star shows WAY too much](#)

[Pratt's father and son moment](#)

[Scientists find brightest object in the universe](#)

[Ghosts found in world's oldest photos?](#)

[Hardwick out for good?](#)

[Lingerie ads win praise](#)

[Instagram trend getting tons of backlash](#)

[Mom's dire slide warning](#)

[Mustang honors WW2 volunteers](#)

[Kate, Meghan team up at Wimbledon](#)

[Duct tape trailer failure](#)

[Billboard causes outrage](#)

[Markle swears by bizarre trick for looking good](#)

[Stealth STD could be next superbug](#)

[Ted Bundy's lawyer tells all](#)

[Previous](#)

[Next](#)

Proud American

[View All](#)

[03:29Video](#)

[Army veteran proves patriotism sells](#)

[U.S.](#)

[New York boy, 3, sings national anthem at baseball game in adorable video](#)

[00:48Video](#)

[Boy vows baseball fans by singing the national anthem](#)

[Health](#)

[Texas sheriff's deputy with 'heart of gold' to donate kidney to man in need](#)

[Cars + Trucks](#)

[Custom Ford Mustang honoring America's WW II Eagle Squadrons to be auctioned for charity](#)

[Tech](#)

<http://www.foxnews.com/>

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

[View All](#)

[13:47hannity](#)

[Hannity: Timing of indictment announcement is suspect](#)

[06:27the ingraham angle](#)

[Interns say Uber driver refused service because of MAGA hats](#)

[05:06tucker carlson tonight](#)

[Behind some Dems' cold feet on abolishing ICE](#)

[04:56fox news @night](#)

[Should Trump press Putin on extraditing indicted Russians?](#)

[09:06the story with martha maccallum](#)

[Graham: Trump going into Putin meeting with NATO stronger](#)

[03:06special report](#)

[Whatever Happened to 'Joe the Plumber'?](#)

Opinion

[View All](#)

[Opinion](#)

[Raise a glass to GOP tax cuts! You can literally drink up the benefits](#)

[Opinion](#)

[Reagan aide on the Trump-Putin Summit: Here is the one crucial thing to focus on before anything else](#)

[Opinion](#)

[Twitter is the latest battlefield between the US and Iran](#)

[Opinion](#)

[Trump-Putin summit could improve US-Russia relations -- And yes, that is a good thing](#)

[Opinion](#)

[Here's what Trump should tell Putin at their summit Monday](#)

[Opinion](#)

[Dan Gainor: Media revel in London anti-Trump protests as obsession with Russia continues](#)

Small Business Spotlight

[View All](#)

[Markets](#)

[The business that's trying to stop the truck driver shortage](#)

[Features](#)

[Teenager becomes millionaire with healthy candy business](#)

[Features](#)

[Golf caddies turn beef jerky into multimillion-dollar business](#)

[Features](#)

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

[Weathering a nasty storm: Secrets from small business](#)

[Features](#)

[Down syndrome entrepreneur builds success out of socks, shatters stereotypes](#)

Full Episodes

[View All](#)

41:00

[Sunday Morning Futures - Sunday, Jun 15](#)

41:02

[The Greg Gutfeld Show - Saturday, Jul 14](#)

40:33

[Justice With Judge Jeanine - Saturday, Jul 14](#)

40:35

[Watters World - Saturday, Jul 14](#)

40:44

[The Journal Editorial Report – Saturday, July 14](#)

41:03

[Fox News @ Night - Friday, July 13](#)

Investing & Stocks

[View All](#)

[Technology](#)

[Apple's betting big on this year's iPhone](#)

[Markets](#)

[The stock-market bull is still on its feet](#)

[Markets](#)

[Commerce's Ross selling stock in response to ethics concerns](#)

[Technology](#)

[Big tech: Here's where, and why, you should invest](#)

[Markets](#)

[Shaky results from Wells Fargo and other banks hit US stocks](#)

[Markets](#)

[Markets Right Now: US stocks open slightly higher](#)

Food + Drink

[View All](#)

[Opinion](#)

[Raise a glass to GOP tax cuts! You can literally drink up the benefits](#)

[Food & Drink](#)

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

'Pioneer Woman' Ree Drummond always keeps this frozen dessert

[Food & Drink](#)

'Papa John' Schnatter claims he was 'pushed' to use n-word

[Food & Drink](#)

Flash mob surprises Chick-fil-A with a capella gospel song

[Food & Drink](#)

Sam's Club offering Polish dogs in response to Costco's news

Travel + Outdoors

[View All](#)

[Lifestyle](#)

Texas fishermen pull 8-foot alligator gar from river: 'She's a beast'

[Cars + Trucks](#)

Paul Ryan says woodchucks 'ate' his SUV -- and he's not alone

[Great Outdoors](#)

Texas woman reels in 7-foot shark: 'Thank God for CrossFit'

[U.S.](#)

Small boat overturns in Hudson River outside New York City; 2 people rescued

[Lifestyle](#)

Moose crashes Alaska couple's wedding, 'obliviously' gallops through ceremony

[Great Outdoors](#)

Great white shark surprises New Jersey fishermen: 'This is crazy'

Autos

[View All](#)

[U.S.](#)

Man, woman steal fire truck, lead police on hours-long chase in California

[Cars + Trucks](#)

Paul Ryan says woodchucks 'ate' his SUV -- and he's not alone

[Cars + Trucks](#)

5 rare and unusual muscle car options you've probably never heard of

[Cars + Trucks](#)

Custom Ford Mustang honoring America's WW II Eagle Squadrons to be auctioned for charity

[Cars + Trucks](#)

Barn-find 1962 Jaguar E-type awakens from 35-year cat nap

[Cars + Trucks](#)

Driver trapped in burning BMW after engine caught fire, brakes failed. Cause unclear

House + Home

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98.310 captures

22 Oct 1996 - 21 Nov 2019



[Why you need a home repair emergency fund](#)

[Real Estate](#)

[Iconic Bel Air mansion owned by Zsa Zsa Gabor up for sale](#)

[Real Estate](#)

[Not sad, but true: Kirk Hammett selling another San Francisco mansion for \\$13M](#)

[Real Estate](#)

[This vacant 1-acre lot in the Bay Area is selling for \\$15M](#)

[Real Estate](#)

[Joanna Gaines reveals 7 easy ways to add 'modern farmhouse' flair to any home](#)

[Cars + Trucks](#)

[150 mph on a lawn mower? Honda Mean Mower V2 takes aim](#)

[Fitness + Well-being](#)

[View All](#)

[Lifestyle](#)

[Aerie's latest inclusive campaign featuring women with disabilities and medical conditions praised online](#)

[Lifestyle](#)

[Times Square billboard for 'appetite suppressant' lollipops slammed by actress, Twitter users](#)

[Lifestyle](#)

[Veteran, 95, jumps off diving board to motivate scared boy](#)

[Lifestyle](#)

[Mississippi couple hosts strangers' wedding inside their home after nuptials are rained out](#)

[Great Outdoors](#)

[Texas woman reels in 7-foot shark: 'Thank God for CrossFit'](#)

[Lifestyle](#)

[Woman who lost 252 pounds rejected on Tinder after sending pics of her excess skin](#)

[Values](#)

[View All](#)

[Opinion](#)

[Go forward in what God is leading you to do in life](#)

[Opinion](#)

[My baby nephew was dying and his mother's response was unforgettable](#)

[Opinion](#)

[Justin, there are 12 things you need to do if you want your marriage to work](#)

[Opinion](#)

[When is it OK to quit church? Here are five reasons to leave](#)

[Opinion](#)

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

[Want to know one surefire way to live longer? It's not what you think](#)

Style + Beauty

[View All](#)

[Lifestyle](#)

[Meghan Markle, Kate Middleton attend Wimbledon in first solo outing](#)

[Lifestyle](#)

[Fashion brand using model with self-harm scars to promote swimwear](#)

[Lifestyle](#)

[Melania Trump praised for J. Mendel gown: 'She looks like a princess'](#)

[Lifestyle](#)

[Aerie's latest inclusive campaign featuring women with disabilities and medical conditions praised online](#)

[Lifestyle](#)

[Times Square billboard for 'appetite suppressant' lollipops slammed by actress, Twitter users](#)

[Lifestyle](#)

[Meghan Markle's trick to looking sculpted involves exercises for your face](#)

Technology

[View All](#)

[Tech](#)

[Microsoft calls for facial recognition technology rules given 'potential for abuse'](#)

[Tech](#)

[Best streaming service: Netflix vs. Vudu vs. Amazon Prime vs. Hulu](#)

[Tech](#)

[Virtual reality developers aim to spark tech boom in the Deep South](#)

[Tech](#)

[Prime Day secrets, best car tech, cool travel apps and more: Tech Q&A](#)

[Politics](#)

[Lisa Page 'cooperative,' 'credible,' lawmakers say after 5-hour closed-door session](#)

[Opinion](#)

[We need to hear more about anti-Trump bias by the FBI and Mueller's team -- House hearing must not be the end](#)

Entertainment

[View All](#)

[Entertainment](#)

[Why Princess Diana's niece doesn't mind being in the spotlight](#)

[Entertainment](#)

[Meghan Markle's dad believes his daughter is 'terrified' adjusting to new life, calls royal family 'outdated'](#)

[Entertainment](#)

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019



98,310 captures

22 Oct 1996 - 21 Nov 2019

About this capture

[Kate Upton announces she is pregnant with her and husband Justin Verlander's first child](#)

[Entertainment](#)

[Dwayne Johnson sees no reason to talk to Tyrese following their feud](#)

[Entertainment](#)

[Nancy Sinatra Sr., first wife of Frank Sinatra, dead at 101](#)

Health

[View All](#)

[Health](#)

[More than 500 fall ill after visiting Tennessee zipline attraction](#)

[Health](#)

[Six North Carolina nurses show off their 'baby bumps' in adorable photo](#)

[Health](#)

[Florida man warns about saltwater poisoning after dog dies days after beach trip](#)

[Health](#)

[Texas sheriff's deputy with 'heart of gold' to donate kidney to man in need](#)

[Health](#)

[What is bacterial meningitis? Signs, symptoms and treatment for the infection](#)

[Health](#)

[Mom split 'like a banana' after horror jet ski accident, report says](#)

Business Leaders

[View All](#)

[Technology](#)

[Apple's betting big on this year's iPhone](#)

[Media](#)

[What's next for Barstool Sports? CEO Erika Nardini weighs in](#)

[Markets](#)

[Papa John's to stop using founder Schnatter's image in marketing campaigns](#)

[Business Leaders](#)

[Northrop Grumman new CEO: What to know about Kathy Warden](#)

[Business Leaders](#)

[What Amazon founder Jeff Bezos plans to charge for a trip to space](#)

[Features](#)

[Meet the Shoe Surgeon, designer of \\$30,000 custom sneakers](#)

Good News

[View All](#)

[U.S.](#)

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

[Indiana patrol officer rescues, later adopts dog after finding it shot in face, covered in bleach](#)

[Health](#)

[Virginia girl, 3, uses FaceTime to save father's life after he suffers stroke](#)

[Sports](#)

[Houston Texans star JJ Watt donates \\$10,000 to family of fallen Wisconsin firefighter](#)

[U.S.](#)

[Pennsylvania mom becomes community hero by feeding hungry children](#)

[U.S.](#)

[Florida firefighters help save man after heart attack, then finish his lawn work](#)

Politics

[View All](#)

[Politics](#)

[HHS departures signal new secretary's 'no-nonsense' approach: report](#)

[03:50Video](#)

[Bolton: Indictment strengthens Trump's hand in Putin summit](#)

[01:51Video](#)

[Trump calls European Union a 'foe' in CBS interview](#)

[Politics](#)

[Trump says 'European Union is a foe' to US ahead of summit with Putin](#)

[06:18Video](#)

[Sen. Orrin Hatch on Democrats' plan to obstruct Kavanaugh](#)

[05:50Video](#)

[Rep. Brady urges Trump to find solution with China on trade](#)

U.S.

[View All](#)

[U.S.](#)

[3 Kansas City, Missouri, cops shot; gunman killed in standoff with police, officials say](#)

[U.S.](#)

[Massachusetts police officer killed after attacked with rock, shot with own gun, officials say](#)

[U.S.](#)

[3 people killed, 22 injured after multi-vehicle crash involving tour bus in New Mexico](#)

[U.S.](#)

[Ski mask, flannel-wearing would-be robber stopped before entering bank, FBI says](#)

[U.S.](#)

[The Latest: Fountain, wing-like benches anchor memorial](#)

[U.S.](#)

<http://www.foxnews.com/>

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

[View All](#)

[World](#)

[French President Macron has plenty of fun at World Cup final](#)

[Politics](#)

[Five key topics that Trump, Putin are expected to discuss in major Helsinki summit](#)

[World](#)

[Fans proud despite Croatia's World Cup final loss to France](#)

[Opinion](#)

[Raise a glass to GOP tax cuts! You can literally drink up the benefits](#)

[World](#)

[Inside Israel's raid to seize nuclear documents in Iran](#)

[Science](#)

[Archaeologists in Egypt discover ancient mummification workshop](#)

Family

[View All](#)

[Entertainment](#)

[Chris Pratt shares touching post about cutting down a tree with his son for Nick Offerman to craft](#)

[Opinion](#)

[Justin, there are 12 things you need to do if you want your marriage to work](#)

[Opinion](#)

[Thai cave rescue: Don't forget man who made ultimate sacrifice so that boys, coach could live](#)

[Opinion](#)

[Ivanka Trump: Paid family leave is an investment in America's families -- It deserves bipartisan support](#)

[Opinion](#)

[How to set your kids up for success -- make sure they are really good at these four things](#)

[Opinion](#)

[Secrets \(good and bad\) no one is telling about motherhood](#)

Sports

[View All](#)

[Sports](#)

[France beats Croatia 4-2 to clinch 2nd FIFA World Cup title](#)

[Sports](#)

[Pussy Riot activists storm field during World Cup match](#)

[Sports](#)

[Novak Djokovic beats Kevin Anderson to win 4th Wimbledon title](#)

[Sports](#)

#4943

<http://www.foxnews.com/>

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98,310 captures

22 Oct 1996 - 21 Nov 2019

[Youthful France to face veteran-laden Croatia for World Cup title](#)

[Sports](#)

[Serena Williams loses to Angelique Kerber in Wimbledon final; dedicates performance to 'all the moms out there'](#)

Crime

[View All](#)

[U.S.](#)

[Colorado cops arrest man with head tattoo who was sought in deadly shooting of woman he knew](#)

[U.S.](#)

[3 Kansas City, Missouri, cops shot; gunman killed in standoff with police, officials say](#)

[U.S.](#)

[Massachusetts police officer killed after attacked with rock, shot with own gun, officials say](#)

[U.S.](#)

[Ski mask, flannel-wearing would-be robber stopped before entering bank, FBI says](#)

[U.S.](#)

[Age-progression video released of missing infant snatched 32 years ago](#)

[U.S.](#)

['Family shootout' in Maine leads to arrests of 64-year-old man, two sons](#)

Fox News Salutes

[View All](#)

[World](#)

[US, North Korea hold talks at DMZ about return of Korean War veterans' remains](#)

[Politics](#)

[Democrat's 'Purple Heart' remark at Strzok hearing spurs veterans to plan protest march](#)

[U.S.](#)

[Paralyzed Florida veteran fights to help others despite physical challenges](#)

[U.S.](#)

[US Navy dedicates destroyer to Sen. John McCain](#)

[World](#)

[Vietnam transfers remains of fallen Americans in ceremony: see the photos](#)

[U.S.](#)

[Burn pits near US military bases in Iraq keep smoldering, as health debate rages](#)

411 Country

[View All](#)

[Entertainment](#)

[Miranda Lambert's rumored beau Evan Felker 'ghosted' wife after health scare, ex claims](#)

[Entertainment](#)

http://www.foxnews.com/

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98.310 captures

22 Oct 1996 - 21 Nov 2019

[Eric Church announces new album 'Desperate Man'](#)

[Entertainment](#)

[Jana Kramer reveals miscarriage struggles before current pregnancy](#)

[Entertainment](#)

[Jason Aldean and wife Brittany expecting baby No. 2](#)

[Entertainment](#)

[Charlie Daniels: The only two things that matter are 'the grace of God and the United States military'](#)

Odd News

[View All](#)

[U.S.](#)

[California man trapped in cement mixer triggers unusual rescue mission](#)

[Science](#)

[Boa constrictor awakens New York man after falling from bedroom ceiling](#)

[U.S.](#)

[Burglar with boxer briefs on head being sought by Texas police](#)

[U.S.](#)

[Armless Miami man charged with stabbing Chicago tourist](#)

[Politics](#)

[Look what she made them do: Rumored Taylor Swift visit leads Capitol Hill interns, media on frantic search](#)

[World](#)

[Giant squirrel art installation in Kazakhstan sparks backlash over \\$67G cost](#)

[U.S.](#)

- [Crime](#)
- [Military](#)
- [Education](#)
- [Terror](#)
- [Immigration](#)
- [Economy](#)
- [Personal Freedoms](#)
- [Fox News Investigates](#)

[World](#)

- [U.N.](#)
- [Conflicts](#)
- [Terrorism](#)
- [Disasters](#)
- [Global Economy](#)
- [Environment](#)
- [Religion](#)
- [Scandals](#)

[Politics](#)

- [Executive](#)
- [Senate](#)
- [House](#)
- [Judiciary](#)
- [Foreign policy](#)
- [Polls](#)
- [Elections](#)

[Entertainment](#)

<http://www.foxnews.com/>

Go

JUN JUL AUG

15

2017 2018 2019

About this capture

98,310 captures

22 Oct 1996 – 21 Nov 2019

- [Style News](#)
- [Entertainment Video](#)

Business

- [Markets](#)
- [Politics](#)
- [Technology](#)
- [Features](#)
- [Business Leaders](#)

Lifestyle

- [Food + Drink](#)
- [Cars + Trucks](#)
- [Travel + Outdoors](#)
- [House + Home](#)
- [Fitness + Well-being](#)
- [Style + Beauty](#)

Science

- [Archaeology](#)
- [Air & Space](#)
- [Planet Earth](#)
- [Wild Nature](#)
- [Natural Science](#)
- [Dinosaurs](#)

Tech

- [Security](#)
- [Innovation](#)
- [Drones](#)
- [Computers](#)
- [Video Games](#)
- [Military Tech](#)

Health

- [Healthy Living](#)
- [Medical Research](#)
- [Mental Health](#)
- [Cancer](#)
- [Heart Health](#)
- [Children's Health](#)

TV

- [Shows](#)
- [Personalities](#)
- [Watch Live](#)
- [Full Episodes](#)
- [Show Clips](#)
- [News Clips](#)

About

- [Careers](#)
- [College Students](#)
- [Fox Around the World](#)
- [Advertise With Us](#)
- [Ad Choices](#)
- [Email Newsroom](#)
- [Media Relations](#)
- [Compliance](#)

Other

- [Fox News Insider](#)
- [Fox News Radio](#)
- [Fox Nation](#)
- [Fox News Go](#)
- [Newsletters](#)
- [Alerts](#)
- [Podcasts](#)
- [Apps & Products](#)
- [Facebook](#)
- [Twitter](#)
- [Flipboard](#)

http://www.foxnews.com/ Go JUN JUL AUG
98.310 captures
22 Oct 1996 - 21 Nov 2019
15
2017 2018 2019 About this capture

[Fox News](#)

- [Terms of Use](#)
- [Privacy Policy](#)
- [Closed Captioning Policy](#)
- [Help](#)

This material may not be published, broadcast, rewritten, or redistributed. ©2018 FOX News Network, LLC. All rights reserved. All market data delayed 20 minutes.

EXHIBIT 12

ESSAY

Open Access to a High-Quality, Impartial, Point-of-Care Medical Summary Would Save Lives: Why Does It Not Exist?

James Heilman^{1,2*}

1 Faculty of Medicine, Department of Emergency Medicine, University of British Columbia, Vancouver, British Columbia, Canada, **2** Wikipedian at Wikipedia

* jmh649@gmail.com



OPEN ACCESS

Citation: Heilman J (2015) Open Access to a High-Quality, Impartial, Point-of-Care Medical Summary Would Save Lives: Why Does It Not Exist? *PLoS Med* 12(8): e1001868. doi:10.1371/journal.pmed.1001868

Published: August 25, 2015

Copyright: © 2015 James Heilman. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: No funding was received for this work.

Competing Interests: I have read the journal's policy and have the following conflicts. I have no financial competing interests. Non-financial conflicts of interest include the following: I am a long-term editor of Wikipedia; I am president of Wiki Project Med Foundation, an NGO that supports the improvement of Wikipedia's medical content; and I speak on the topic of Wikipedia and medicine at conferences, for which I do not receive honorariums, but I have previously received some travel support from the NIH, UCSF, WMF, and Cochrane Collaboration. In July of 2015 I became a Board Member of the Wikimedia Foundation, the not-for-profit that runs Wikipedia.

Abbreviations: BY, by attribution; CC, creative commons; NC, non-commercial; ND, no-derivatives;

Summary Points

- Currently no open access point-of-care (POC) medical summary aimed at a professional audience exists.
- Some nonprofit and multiple professional, for-profit POC medical summaries are frequently accessed by clinicians and policymakers.
- Efforts to create open access POC summaries have been stymied by the difficulty of attracting high-quality contributors.
- The open access medical publishing community can create this resource with engaged donors, crowd-sourcing, and technology.

The Current State

Over the past decade, the world of scientific journal publishing has been transformed by open access (OA) to information. In the strict sense, open access refers to the ability for others not only to view but also to build upon and distribute a work as long as attribution of the author(s) is provided [1]. *PLOS ONE* became the world's single largest journal (by number of articles) only four years after it was founded and has since increased in volume nearly 5-fold [2,3]. Currently there are nearly 10,000 journals listed by the Directory of Open Access Journals; as of May 2015, more than 1.9 million articles [4]. About a quarter of these were related to medicine. In 2011, of all scholarly articles published, 17% were OA [5], and in the biomedical fields the proportion of freely available articles (both OA and “free” access) passed the 50% mark in 2010 [6].

The huge increase in access to scientific knowledge has been chiefly of benefit to those researchers who have the time to search the literature. It is less helpful for working health care providers, since masses of literature do not lend themselves to reliably and promptly answering questions. A recent estimate that 85% of all medical research is wasted is based on waste in the research process and in publishing the research itself [7]. This problem of waste is compounded when medical knowledge exists but health care providers have to make decisions without it, a

NGO, non-governmental organization; OA, open access; POC, point of care; SA, share-alike.

Provenance: Commissioned; externally peer reviewed

phenomenon that has not been properly examined. One survey found that 80% of primary treatments given at a university hospital in the United Kingdom had some evidence to support their use [8]. This study, however, did not look at whether better evidence existed or how well the evidence matched the treatment decision or patients' values, nor does it likely represent typical practices. Others have claimed that the typical rate is closer to 10% to 20% [8].

Point-of-care (POC) medical summaries exist to help physicians make decisions in ways not well served by traditional publishing formats like journal articles and textbooks. POC summaries have been described as "web-based medical compendia specifically designed to deliver pre-digested, rapidly accessible, comprehensive, and periodically updated information on a given topic to health care professionals" [9]. Unlike PubMed, Cochrane, and other resources, POC tools are designed to be quick to search and navigate and aim to help health care providers solve emerging clinical questions. As such, they are collections of "review articles" browsable in real time to aid in patient care.

What Is Already Available?

A number of point-of-care medical summaries are available. Some are open access but not aimed at a professional audience, while those that are aimed at a professional audience are not open access. A few of the non-open access, web-based evidence summaries, while freely accessible, are funded by advertising revenue, making them vulnerable to distortion by financial interests [10]. Others, such as National Health Service's Clinical Knowledge Summaries, are only available in Great Britain. The most popular tools aimed at professionals are subscription-only, provided by several wealthy, for-profit publishers: UpToDate and DynaMed are the most successful and best known.

Health care providers and institutions are willing to pay to access high-quality point-of-care medical summaries. Personal subscriptions for DynaMed and UpToDate are several hundred dollars for a physician and substantially more for institutional subscriptions. UpToDate company data states that its entries are the result of over 6,000 authors, editors, and peer reviewers. It reports that it is used by over "1 million clinicians in 174 countries and [that] almost 90% of academic medical centers in the United States rely on UpToDate to provide the best care." The word "rely" is a curious one, but usage statistics support the idea that health care providers find UpToDate helpful, with more than 23 million topic views logged each month [11]. The company advertises research showing that its use is associated with "...reduced length of stay, lower risk-adjusted mortality rates, and better quality performance" [12]. Based on visits, UpToDate was within the top 50 health sites on the Internet in March of 2015 [13].

While the value of these tools cannot be assessed by observation or anecdote alone, it is clear that clinicians seek out, use, and personally pay for them. The wider world is also interested in health care content. Google states that one in twenty searches is for health care information [14]. Additionally, Wikipedia's popularity among both physicians and the public is a further indication of global demand [15].

What Hasn't Succeeded?

Given the rise of OA scientific publishing, why have we not seen similar developments making the world's medical knowledge freely available in a constantly updated and intuitively accessible format ready at the point of care? There have been attempts. But they serve not only as examples that such databases are desired but also that creating them is difficult.

In 2009, Medpedia was formed with venture capital, and their website stated they had the support of top universities, including Harvard Medical School, Stanford School of Medicine, and the University of California Berkeley School of Public Health [16]. Medpedia never caught

on, and with few readers or contributors, it ceased in 2013. In 2011, WikEmerg.ca was created. It initially paid physicians to write content, but experienced trouble getting engaged contributors even when money was attached. Efforts to create broader professional encyclopedias have also failed. One of the first was Nupedia in 2000, which grew slowly under Jimmy Wales and Larry Sanger. Wales and Sanger decided to launch a website with looser contribution requirements and, in 2001, created Wikipedia. Nupedia faded, but Sanger gave the model a second try in 2007 with the launch of Citizendium. As of September 2014, it had less than 160 high-quality articles and struggled to raise enough money to stay online [17]. Even Google has tried and failed. In 2008, they started a supposed “Wikipedia killer” called Google Knol [18]. The site closed in 2012. Google Knol’s failure has been attributed to several factors, including low numbers of contributors, poor reward mechanisms, and low potential for collaboration. It rapidly filled with plagiarism and poor-quality content [19].

Wikipedia

In medicine, the one web-based, crowdsourced project that seems to have had a degree of success is Wikipedia [15]. Ironically, the platform is geared towards the general population and not health care providers. Furthermore, Wikipedia does not intend readers to use the site to treat themselves or others. The quality of Wikipedia entries is a topic of vigorous debate, and even with its semiformal internal peer review process, the small size of the active medical community means fewer than 1% of medical articles have passed review. Yet medical information on Wikipedia is widely accessed. About 5 billion page views were logged for its medical content in 2013. Of medical students, 92% to 94% use it, and 50% to 70% of practicing physicians report also doing so [15,20,21]. A 2009 survey of UK junior physicians found that 70% use Wikipedia on a weekly basis, making it the most used content provider, ahead of eMedicine.com, which was used by nearly 50% [22]. These numbers, as with those for UpToDate, suggest a widespread desire and need for point-of-care clinical information.

Creating an Open Access Point-of-Care Summary Content

An open access professional database would need open access professional content. Such a database could be created “de novo,” it could be developed from currently available open content, or it could be constructed by releasing what is currently closed-source content under open licenses. Even if a benefactor with deep pockets were willing to buy UpToDate or DynaMed and make its content freely available as a starting point, mechanisms to regularly update the content remain essential.

Creating entirely new content is an enormous undertaking, but with the support of OA publishing and the wider medical community, it may be possible. UpToDate has around 10,000 topics [9], while eMedicine is estimated to contain 7,000 [23]. Likewise, Wikipedia contains about 7,000 disease- or symptom-related articles. As previously mentioned, there are about 2,500 medical OA publishers. If each was to commission a single broad-topic review each year, they would quickly match the scope of content contained within these commercial sites.

There is a great deal of existing open access content that could be adapted to be more professional and clinically useful. This includes content from Wikipedia, the US Centers for Disease Control and Prevention, the Food and Drug Administration, and the National Institutes of Health (NIH). The NIH, for example, has open access content spread across at least 27 different institutional websites. Even finding the correct website is, therefore, a challenge when searching for the topic one is looking for. Additional material could come from other organizations if they could be convinced to release content more openly; these might include the World

Health Organization, National Health Services in the United Kingdom, and Health Canada. The Bulletin of the World Health Organization began using an open license in July of 2014. Medication-related information could possibly come from organizations such as the American Society of Health-System Pharmacists or the Royal Pharmaceutical Society, which already produce the American Society of Health-System Pharmacists Drug Information and British National Formulary, respectively.

A team of editors would then give the final product a consistent style, find appropriate peer review, and verify quality. How, though, should this team and the website by which the information was delivered be paid for?

Funding Model

Expenses for a medical project such as this relate not just to platform development but also content creation, editing, production, and promotion. Recruiting quality content has proved a stumbling block for some previous efforts despite offering payment for written work. Unpaid content creation draws upon the established practice of many scientific journals; editors commonly expect experts to write review articles for free, with the joint rewards of increased visibility for their ideas and a new publication on their professional record. For an open access POC medical summary to attract quality content, it must likewise confer prestige or career advancement to its authors. Most open access journals, including *PLOS Medicine*, do not contain literature-review-type articles (not to be confused with systematic reviews). Expecting authors to pay to write articles, as in a current open model for primary research, would not be realistic for reviews unless publication, as with research, also lead to career advancement.

Wikipedia has a surprisingly small community of contributors, perhaps partly because it lacks prestige among the academic community. If contributing to a global knowledge database came to be valued, in terms of prestige and career progression, contributors would compete for the right to help. Once the reputation of such a database became established, it would be self-reinforcing, encouraging ever more competition to contribute, with resulting increases in quality and prestige for those involved. There would also be something honourable about spending part of one's career on such an effort.

Wikipedia's funding model relies almost entirely on donations from readers. This is relatively effective, even given the small percentage who contribute, as nearly half a billion individuals use the site each month. About 2 million people donated an average of US\$15, mostly during a one-month period last year. Still, Wikipedia runs on a low budget of less than US\$50 million a year, as compared to annual revenues of US\$4.7 billion for the parent company of UpToDate [11] and US\$2 billion for the parent company of DynaMed [23]. The Wikimedia Foundation, the non-governmental organization (NGO) behind Wikipedia, developed slowly, and it is only in the last few years that they surpassed 100 employees and US\$10 million [24]. To seriously undertake the establishment of a global clinical knowledge database, of sufficient quality to become self-sustaining through acquiring status and accepted utility, would require inputs of tens of millions of dollars a year. Arguably it might also be an incredibly cost-effective way of improving global health, simply by making sure the best medical knowledge was freely available to those who needed it.

Copyright

The NIH currently pays for licenses to use some material for their medical encyclopedias. Even though hosted on a .gov website, it is not in the public domain or under an open license. The content is basic, containing no inline references, and is too simple to be useful to a professional.

The World Health Organization (WHO) and National Health Services (NHS) have much high-quality material that can be freely viewed online, but very little of it is under an open license. The WHO material, additionally is often not in a format that is easily viewed, with much of their content in large, image-laden PDFs that are slow to load even for those with excellent internet connections. While the WHO's policy on open access is now supporting greater availability of the material it produces, it has not yet gone so far as to make all of its content available under an open access license, such that others can build upon it. Additionally the organization occasionally uses the term "open access" when they appear to mean "free access" [25]. One would hope that it would be possible to persuade these organizations to embrace true open access on the grounds that this release of knowledge would be consistent with their aims of advancing health.

The US federal government has been releasing much of what they produce into the public domain since 1895 [26]. The benefits to science and health of this position have been much greater than any harms. This is the reason, for example, we have access to the complete human genome. There is little evidence that others have taken their work, altered it, and claimed that it is still the position of the US government. United Nations Educational, Scientific, and Cultural Organization (UNESCO) began using open access licenses in July of 2013, as they see this as "key to the advancement of innovative solutions for the challenges of international development" [27]. This applies to all content published after July 2013, with some older content also being relicensed.

Problematic among those who do use open licenses is that there is not just one creative commons (CC) or open license, but many exist and they are often incompatible or only compatible in one direction. Qualifiers of open licenses include "by attribution" (BY), "share-alike" (SA), "non-commercial" (NC), and "no-derivatives" (ND). Many NGOs create content under CC BY SA NC licenses that are not usable within a CC BY or CC BY SA source. If such acronyms seem confusingly bad reasons for not sharing knowledge that could improve global health, so they should. Efforts at Wikipedia to improve reliability included efforts to work more closely with PLOS and BMJ Open, but these efforts faced copyright issues. A common license for open access journals is CC BY, while Wikipedia's content is under an incompatible CC BY SA. Thus, Wikipedia content cannot be published by these journals under their usual license. Neither PLOS, BMJ Open, nor Wikipedia were able to find a way to align their licenses. These licensing issues make it difficult for those who are working towards the same goal of "health information for all" to collaborate. The NC license faces the additional issue that what defines a non-commercial use is not clear. For example, while many view education as non-commercial, students pay to attend university, and professors make a living teaching and doing research.

Global Access

Getting access is key. Simply putting content up on the Internet is not enough—many people, including many health care providers in low- and middle-income countries, do not have easy or inexpensive access to the Internet. Most people, however, do have access to cell phones. The Wikimedia Foundation has agreements with cell phone companies to provide Wikipedia without data charges in low- and middle-income countries. As of May 2015, 400 million people in low- and middle-income countries have free access to Wikipedia without data charges through these programs. This suggests a model for widely disseminating a clinical knowledge database. Loading health content on phone systems before they are shipped may be another option and is one Wikipedia is currently pursuing.

Looking Forward

The success of PubMed Central and the many open access journals in making medical knowledge openly available to researchers has been immense. Nothing similar has yet been done for clinicians. Our immediate goal should be to provide accessible, high-quality medical information that is easily searchable and usable at the point of care ([Box 1](#)). Using for-profit POC medical summaries as a yardstick, this would involve the putting together of 5,000 to 10,000 overview articles. To have global impact, this content would need to be available in languages comprehensible by the majority of the world's health care providers. It would need to be updated on a regular basis and available in multiple formats, including on the web, as e-books, and in spoken form. It would also need to be freely readable, including without data charges in the developing world.

Box 1. Options

Costs low, risk minimal, control limited

Established medical publishers work to engage clinicians, academics, and government agencies in an effort to generate high-quality, open access POC summaries. Open access summaries could initially be developed via commissions from high-impact journals. Once a database of summaries is established, prestige will reinforce content contribution. Experts can follow their conscientious desire to promote global health while benefiting professionally.

Open access journals would need to agree on a consistent style and compatible license. Funders of research such as the NIH and NHS could put money towards these efforts and benefit from reuse of the materials on their own websites.

This is already happening to a limited extent in a slightly different format, with the first Wikipedia article formally published following peer review in the journal *Open Medicine* in October of 2014 [28]. Cancer Research UK and ECGWiki have released large number of images under a more open license and organizations including the Cochrane Collaboration, UCSF School of Medicine, Cancer Research UK and the National Institutes of Health are working to improve Wikipedia content.

Costs moderate, risk moderate, control moderate

A medium-sized not-for-profit could be set up and funded through grants, donations, and goodwill. Membership would come with expectation of openness. This would hopefully push some organizations to reconsider their position and become more serious about the importance of open licensing.

Create and manage a MediaWiki website editable only by approved users. Build up a collaboration of partners. Develop a manual of style and outlines. Petition partners to fill in portions with stipends and authorship as rewards. Use both internal and external peer review. Encourage content from Wikipedia and inclusion into Wikipedia to expand audience. Hire programmers to manage automated scripts and develop the underlying software to better fit the needs of the site.

Costs significant, risk large, control greater

Build from the ground up, create a new website and software, hire authors, hire translators, convince cell phone companies to allow free use, and raise capital. Would require convincing major funders to invest large amount yearly and in perpetuity. Would require hundreds of employees to manage all aspect of a large publishing and distribution house.

Such a database is not a small goal, but it is a worthy one. That medical decisions are made without timely access to the existing evidence base, and that this occurs in a world which recognises the importance of open access to scientific knowledge, should not be allowed to continue. An opportunity exists for governments, NGOs, IGOs, and philanthropists to reduce the global burden of disease simply by working together to make existing medical knowledge easily, freely, and usefully available.

Author Contributions

Wrote the first draft of the manuscript: JH. Contributed to the writing of the manuscript: JH. Agree with the manuscript's results and conclusions: JH. ICMJE criteria for authorship read and met: JH.

References

1. Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. <http://openaccess.mpg.de/Berlin-Declaration>. Accessed: 2014-11-20. (Archived by WebCite at <http://www.webcitation.org/6UEkjbvbcR>)
2. Konkiel, Stacy; PLoS ONE: Five Years, Many Milestones. PLOS Blogs. <http://blogs.plos.org/everyone/2011/12/20/plos-one-five-years-many-milestones/>. Accessed: 2015-07-09. (Archived by WebCite at <http://www.webcitation.org/6ZtzKsfQU>)
3. Wikipedia contributors. PLOS ONE. Wikipedia, The Free Encyclopedia. July 4, 2015, 17:40 UTC. https://en.wikipedia.org/w/index.php?title=PLOS_ONE&oldid=669943074. Accessed July 9, 2015.
4. DOAJ. Directory of Open Access Journals (DOAJ). DOAJ. 2014-09-02. <http://doaj.org/>. Accessed: 2014-09-02. (Archived by WebCite® at <http://www.webcitation.org/6SHZY1PVh>)
5. Laakso M; Björk BC (22 October 2012). Anatomy of open access publishing: a study of longitudinal development and internal structure. BMC medicine 10: 124.
6. Kurata K; Morioka T; Yokoi K; Matsubayashi M (2013). Remarkable growth of open access in the biomedical field: analysis of PubMed articles from 2006 to 2010. PLoS ONE 8 (5): e60925.
7. Glasziou P (May 2014). The role of open access in reducing waste in medical research. PLoS Med 11 (5): e1001651.
8. Ellis J; Mulligan I; Rowe J; Sackett DL (12 August 1995). Inpatient general medicine is evidence based. A-Team, Nuffield Department of Clinical Medicine. Lancet 346 (8972): 407–10.
9. Banzi R; Liberati A; Moschetti I; Tagliabue L; Moja L (7 July 2010). A review of online evidence-based practice point-of-care information summary providers. Journal of medical Internet research 12 (3): e26.
10. Heffernan, Virginia. A Prescription for Fear. The New York Times. 2014-09-02. http://www.nytimes.com/2011/02/06/magazine/06FOB-Medium-t.html?_r=1. Accessed: 2014-09-02. (Archived by WebCite at <http://www.webcitation.org/6SHYzl2TD>)
11. UpToDate. About Us: UpToDate. <http://www.uptodate.com/home/about-us>. Accessed: 2015-07-09. (Archived by WebCite at <http://www.webcitation.org/6Zu0B2uHi>)
12. Isaac T; Zheng J; Jha A (February 2012). Use of UpToDate and outcomes in US hospitals. Journal of hospital medicine: an official publication of the Society of Hospital Medicine 7 (2): 85–90.
13. Website Ranking Top 50 sites for Health. <http://www.similarweb.com/category/health>. Accessed: 2015-03-01. (Archived by WebCite at <http://www.webcitation.org/6WiLHNoVY>)

14. Ramaswami, Prem. A remedy for your health-related questions: health info in the Knowledge Graph. <http://googleblog.blogspot.jp/2015/02/health-info-knowledge-graph.html>. Accessed: 2015-03-01. (Archived by WebCite at <http://www.webcitation.org/6WiLN36Lr>)
15. Heilman JM, Kemmann E, Bonert M, Chatterjee A, Ragar B et al. (2011). Wikipedia: a key tool for global public health promotion. *J. Med. Internet Res.* 13 (1): e14.
16. Silverman, Ed. Sites strive to make medical information accessible. <http://healthjournalism.org/blog/tag/medpedia/>. Accessed: 2014-11-20. (Archived by WebCite at <http://www.webcitation.org/6UEjkqpAf>)
17. RationalWiki. Citizendium: Fiscal insolvency: the \$65,000 question. http://rationalwiki.org/w/index.php?title=Citizendium&oldid=1376857#Fiscal_insolvency:_the_2465.2C000_question. Accessed: 2014-11-20. (Archived by WebCite at <http://www.webcitation.org/6UEkCD1Lw>)
18. Karch, Marziah. Annotum Definition. http://google.about.com/od/experiment_graveyard/g/Annotum-Definition.htm. Accessed: 2014-09-02. (Archived by WebCite at <http://www.webcitation.org/6SHUYF1Af>)
19. Bailey, Jonathan. Why Google Knol Has Failed. <https://www.plagiarismtoday.com/2008/09/23/why-google-knol-has-failed/>. Accessed: 2014-11-20.
20. Allahwala UK, Nadkarni A, Sebaratnam DF. Wikipedia use amongst medical students—new insights into the digital revolution. *Medical teacher* 2013; 35 (4): 337.
21. Adomi E, Adigwe S. Use and Perception of Wikipedia among Medical Students in a Nigerian University. *International Journal of Digital Literacy and Digital Competence* 2014 Apr; 5(2):1–11
22. Hughes B, Joshi I, Lemonde H, Wareham J. Junior physician's use of Web 2.0 for information seeking and medical education: a qualitative study. *Int J Med Inform* 2009 Oct; 78(10):645–655. doi: [10.1016/j.ijmedinf.2009.04.008](https://doi.org/10.1016/j.ijmedinf.2009.04.008) PMID: [19501017](https://pubmed.ncbi.nlm.nih.gov/19501017/)
23. Dean. Point-of-care tools in medicine. http://hlwiki.slais.ubc.ca/index.php/Point-of-care_tools_in_medicine. Accessed: 2014-09-02. (Archived by WebCite at <http://www.webcitation.org/6SHYIHxd3>)
24. Wikipedians. Wikimedia Foundation. Wikipedia. https://en.wikipedia.org/?title=Wikimedia_Foundation. Accessed: 2015-06-23. (Archived by WebCite at <http://www.webcitation.org/6ZUwvVhab>)
25. WHO policy on open access. World Health Organization. <http://www.who.int/about/policy/en/>. Accessed: 2015-07-11. (Archived by WebCite at <http://www.webcitation.org/6ZwqHbeV0>)
26. Copyright in Government Publications, in: Copyright Law Revision: Studies Prepared for the Subcomm. on Patents, Trademarks, and Copyrights of the Senate Comm. on the Judiciary, 86th Cong., 2d Sess. 29–30 (Comm. Print 1961)(Study 33), pp. 27.
27. UNESCO. Open Access Policy concerning UNESCO publications. http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ERI/pdf/oa_policy_rev2.pdf. Accessed: 2015-06-23. (Archived by WebCite at <http://www.webcitation.org/6ZUrfQhnx>)
28. Hesp-News. Dengue fever: a Wikipedia clinical review. <http://hesp-news.org/2014/10/02/dengue-fever-a-wikipedia-clinical-review/>. Accessed: 2014-11-20. (Archived by WebCite at <http://www.webcitation.org/6UEkTcNBf>)

EXHIBIT 13



Who should be my Facebook partner? Analysis of the relationship between Hungarian large-scale Facebook pages

Pálma Szolnoki¹, Márton Ilyés¹, László Gulyás², and George Kampis^{2,3}

¹ Corvinus University of Budapest, Hungary

szolnokipalma,ilyes.marton@gmail.com

² Eötvös Loránd University, Budapest, Hungary

gulyas.zana.laszlo, kampis.george@gmail.com

³ ITMO University, St. Petersburg, Russia

Abstract

We analyze the relationship between large-scale Hungarian Facebook pages, where connection is based on a common active fan base. For this purpose, we built a network database (121 nodes and 655 edges) where the nodes are Facebook pages and the edges indicate if a given Facebook page has a common active fan base with another Facebook page. According to the results, the primary organizing principle in this relationship is the topic, while other factors such as ownership background and political views could not be identified as ordering principles. Within large topics, gender and social background are identified as sub-ordering principles. A further result is that the identified clusters could provide useful input for Facebook page operators in finding valuable link-exchange partners.

Keywords:

1 Introduction

The importance of Facebook has grown over the years. In 2015 among the top American news sites the traffic coming from Facebook outgrew that of Google's [4]. Similarly to international trends, the importance of Facebook as a referral site also grew among Hungarian portals. There are various Hungarian webpages (news sites, magazines, etc) that gain their revenue almost entirely based on the number of visitors, therefore they are extensively dependent on referral sites such as Facebook. There are pages where the 90% of the viewers come from Facebook¹, and even in the case of the most popular Hungarian news site, Index.hu, about 13% of its visitors come from Facebook. Recognizing the importance of Facebook, the Hungarian portals

¹as for example one of the most popular portal, Mindenegybenblog.hu:
<https://www.similarweb.com/website/mindenegybenblog.hu#overview>

Who should be my Facebook partner? ...

Szolnoki, Gulyás and Kampis

thus devote more resources to try to understand the behavior of Facebook users. For instance, similarly to the international sites [5] they started to look for partners with whom they can exchange link-posts. This way they hope to find new visitors/likers that would be interested in their sites. Such cooperation could be seen for example on the Facebook pages of Mindenegyben blog and Kiskegyed².

It could help in finding valuable exchange partners if the sites knew which topics go along well with other topics, e.g. if a user is active in one topic on Facebook, then typically what other topics she/he would be also active/interested in. This is important because partnering is unlikely within a topic, since competitors do not want to help each other.

In this article we try to identify these kinds of relationships within the Hungarian Facebook pages, where the basis of the relationship is the similarity of active fan base on FB pages. In the first section we introduce the research question. In the second section we describe the network, and this is followed by the network analysis in the 3rd section and a summary of the results in the 4th section.

For the analysis, we selected Hungary because Hungarian Facebook pages constitute a sizeable yet easy-to-handle set and the authors are Hungarian thus there is no language barrier.

2 Research Question

First we summarize the results of the most recent survey about the behavior of the Hungarian Facebook users, and based on these, we define a hypothesis.

2.1 Background

The Association of Hungarian Content Providers (MTE) hired NRC NetPanel to conduct a representative online survey among the 18-49 year old Hungarian Facebook users using a sample of 500 persons [1]. The survey was taken in November 2015. The results were published in February 2016. From the perspective of this article, the followings were the most important results.

Within the investigated age group 82% of users are daily users ([1], slide 8), and on daily basis they tend to spend almost one and a half hour (86 minutes) on the average just scrolling Facebook ([1], slide 10). Based on the answers of the users, the study defined three user types:

- 35% of the analyzed users are Heavy users, being active on Facebook more than one and a half hour daily.
- 34% of the users are Medium users, scrolling daily 30-90 minutes, and
- 30% are Light users scrolling a maximum of 30 minutes daily.

”Among the heavy users women, the lowly educated and people under 30 are overrepresented.” ([1], slide 10)

The NRC study distinguished between passive and active activities on Facebook. The former category consists of activities that cannot be detected by other FB users, such as newsfeed

²see their Facebook page posts:

- On the page of Kiskegyed a Mindenegybenblog.hu post:
<https://www.facebook.com/138229586210610/posts/1184365694930322>
- On the page of Mindenegybenblog.hu a Kiskegyed post:
<https://www.facebook.com/Mindenegybenblog/posts/1130214633718832>

Who should be my Facebook partner? ...

Szolnoki, Gulyás and Kampis

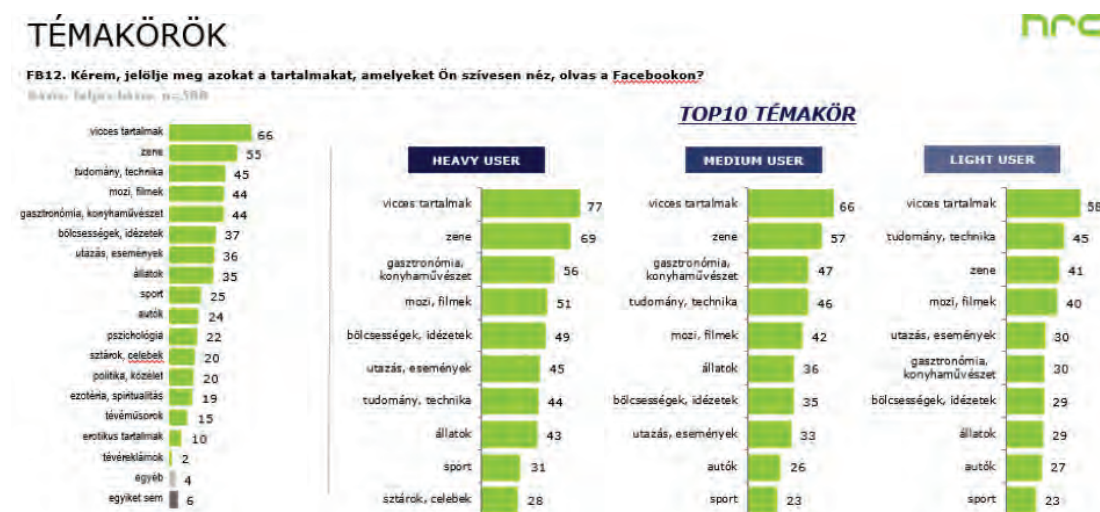


Figure 1: Topics ranked according to popularity among the Hungarian Facebook users in the age group 18-49 years. (In Hungarian. Vicces=funny, zene=music, tudomány=science, mozi=movies and so on.

scrolling, link clicks, etc. The latter consist of activities such as liking, sharing, commenting, posting, uploading pictures, etc. which can be seen by others in FB. About 27% of the sample state that they take active actions frequently, this is distributed over the user types as follows: ca. 40% of Heavy users, ca. 25% of Medium users, and ca. 13% of Light users ([1], slide 18).

In the circle of the analyzed users, the most popular Facebook pages were those that were created exclusively for Facebook (average 49%, Heavy users 55%), these are followed by the FB pages of the news sites and magazines (43%, Heavy users 54%), then come the blogs (30%, Heavy users 38%) then the FB pages of TV-radio channels and their shows (19-19%, Heavy users 32-31%). Public figures and music bands have a relatively low fan base (1-1%, Heavy users 1-1%). ([1], slide 26)

Within those FB users who follow portals and blogs, the 87-77-80% of Heavy-Medium-Light users frequently like their content, 71-60-55% share their content, and 53-31-28% comments frequently on the content. ([1], slide 21.)

The plot of Figure 1 shows which topics are popular on Facebook on average and within user groups. On average, the most popular content among Hungarian FB users are funny posts (66%), these are followed by music (55%), then science and technology (45%) along with movies (44%) and gastro pages (44%). A little less popular are quotes (37%), traveling (36%) and animals (35%), respectively. These are followed by sports (24%) along with cars (25%), then psychology (22%) and celebrities (20%) and after that politics (20%). This ranking differs a bit according to the various user types. In case of Heavy and Medium users the gastro pages are ranked 3rd and for Heavy users the quotes are ranked higher while science and technology are ranked lower. ([1], slide 25)

2.2 The Research Hypothesis

Available Facebook data describes the followers of FB pages. It can be seen what percentage of the followers of a Facebook page are actually active, i.e. by the likes, shares and comments

Who should be my Facebook partner? ...

Szolnoki, Gulyás and Kampis

to the page's posts, and it can also be seen on which other FB pages this active fan base also active. This latter relationship forms the basis of our network.

Although the MTE survey (MTE, 2016) was restricted to FB users aged 18-49 years while our data refers to all FB users, still this is the most comprehensive survey on Hungarian FB users. Therefore we will base our upfront hypothesis on the MTE findings. On the database thus built, we investigate what is the primary organizing principle of the Hungarian Facebook users' activity: is it the topic that is determinative, or can there be other organizing principles such as gender, age group, political views, the ownership structure of the pages, etc. Furthermore, should the topic turn out to be the most determinative in the users' activity, the the question offers itself, which topics connect more to each other, and within this connection what other organizing principles may play a role.

Since the database is based on the active users' activity, therefore for our hypothesis we use those findings of the MTE study which describe the active users. About 1.5 times more Heavy users take active actions than do Medium users and again 3 times more Heavy users than Light users. Therefore the content ranking of Heavy users is the most relevant for our analysis. Furthermore, since within Heavy users the groups of women, the lowly educated and those under 30 are overrepresented, therefore we predict that, within the respective topics, those FB pages that aim at these groups would be determinative. Therefore our hypothesis is that jokes, music, gastro, movies, and quotes are the topics the would be determinative in the network, i.e. forming independent clusters, and if these topics connect within a cluster, then those connections would reflect the background of Heavy users, so for example women-tabloid-quotes could form such a cluster, etc.

3 The database

The database is based on the information provided by the Fanpage Karma (FPK) software [2]. FPK provides structured information on all Facebook pages for the previous six months. Besides basic data like fan base size, it also provides less visible information on FB pages, like what percentage of the fan base is active, who are these active users, and on which FB pages these active users are also active. This latter forms the basis of our study network, where nodes are the Facebook pages, while directed edges represent a common active fan base. Edges are directed, as the relationship as we will see is usually not symmetric, which could be attributable to the large differences in fan base size.

The relationship is published by FPK for the last 28-day interval, the database represents the connections in May-June 2016. We built the database with the snowball method. We started out from the well-known Nosalty Hungarian gastro site's Facebook page and looked up on which FB pages are its active fans also active, and went along those pages and so on. In doing so, we made two restrictive criteria for going forward:

- only Hungarian FB pages were considered
- only pages that have a fan base of 50,000 or more users were counted

The first restriction is straightforward, as we want to investigate the relationships between Hungarian FB pages. Therefore international or foreign sites would only distract the analysis. There were very few cases when we had to leave out a page for this reason, and those cases were typically in the topics of sports and video games. We applied the second restriction because we were interested in truly influential Hungarian FB pages. The 50,000 limit is arbitrary, set as a thumb rule during the application of the snowball method. Finally, we closed the database

Who should be my Facebook partner? ...

Szolnoki, Gulyás and Kampis

also arbitrarily after 121 nodes. There were 5 FB pages where we could have continued with the snowball method further and these are indicated in the database.

3.1 The Data

We built a network with 121 nodes and 655 edges. To the nodes, i.e. Facebook pages we collected further data as attributes:

- Fan base: the size of the page, i.e. how many total followers the FB page has.
- Engagement: This is an FPK indicator, it shows how active is the fan base. Engagement is a very good proxy about which sites are really vital and which were only inflated once with some kind of a sweepstake, but in reality have no real fan base.
 - Calculation: the sum of daily likes, shares and comments is divided by the total fan base (FPK Engagement, 2016)
- Ownership: we looked up whether the given FB page is part of a larger holding, or whether is it an individual player. For this we checked the FB Page About section, the webpages' impressum and media offer sections. (This category was only filled out in 42 cases, the rest we considered as individual pages.)
- Type: This category describes the purpose of the owner with the FB page. Whether it is a page created solely for Facebook or it is related to a webpage. Within the latter there are various types: there are those where the Facebook page is set up mainly in order to drive traffic to the portal (deflector), news sites are like this for example; there are those cases when the FB page is for branding, e.g. Pickwick; finally within this type there are the political parties and public figures which use Facebook pages also for branding and mobilizing. Within the deflector sites there are those which we named collectors. This kind of Facebook Page is created besides the official Facebook page of a webpage, also for driving traffic for the webpage. The Konyhatündérek-Receptneked Facebook page is for example a collector FB page for Receptneked ("Recipes for you"), and the Imádok sütni ("I love to bake") FB page is a collector for Nosalty. Finally there are the FB pages of public figures.
- Topic: This arbitrary categorization describes the topic of the FB page. Main categories are news, gastro, politics, video games, ...
- Subtopic: Within some topics we made further differentiations, for example within the music topic we differentiated radio show, singer, showmen, etc.

4 Network Analysis

As mentioned, the database was built with the snowball method, where from each node 1-8 edges were directed towards other edges, and then it was closed arbitrarily. This method already determines some of the network characteristics:

- the network is always at least weakly connected;
- the outdegrees vary between 1-8 except for the (five) closing nodes, where the outdegrees are 0.

For the network analysis, we used the Gephi package.

Who should be my Facebook partner? ...

Szolnoki, Gulyás and Kampis

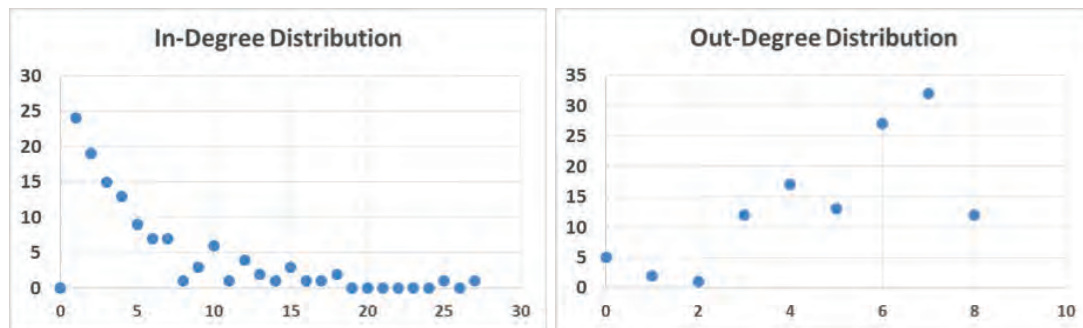


Figure 2: The degree distributions.

4.1 Basic network characteristics

The network density is found to be low, 0.039, which is a usual phenomenon for networks of this size. The distribution of indegree values (Figure 2a) provides an interesting insight on the relationship of the Hungarian Facebook pages. There are two pages with unusually high indegree values: 27 - Blikk and 25 - Startlap. Now, Blikk is the largest tabloid newspaper in Hungary, and Startlap is one of the main news-link collector sites operated by a large media group with many affiliate sites.

There are 7 other Facebook pages with indegree values over 15, and six of these are operated by major news sites and one is a Facebook page of a women's magazine. About 90% of all the nodes has an indegree value lower than 15. This leads to a distribution which is similar to powerlaw distribution: the majority of sites has only a few relationships pointing towards them and only a few have many incoming relationships. Presumably if the network contained all of the Hungarian Facebook pages, the distribution of indegree values would be even closer to a powerlaw-like distribution.

The diameter of the network is 18, and average path length is 5.8. These values could be considered relatively large, compared to the size of the network. Presumably this is due to the directedness of the network.

The network is weakly connected, and it has 14 strongly connected components. Exactly 11 of these have only one element:

- 5 pages, these are the ones where the network snowballing was closed arbitrarily,
- 3 pages have very low engagement values which could mean that these are not vivid sites, therefore the absence of relationship with other sites is understandable, and there are
- 3 pages, where it is not straightforward, why they are not connected to other sites more strongly, and presumably if the network would be larger, these pages would not stand alone.

There is one component with two nodes, these are two women's sites targeting young women and have a strong relationship: same columnist, many link exchanges. These two are standing out from the crowd of women magazine FB pages, as the rest are targeted towards older readers.

There is one component with 11 elements. These are all brand pages with similar content from the food-furnishing and drug store businesses.

The largest component has 97 elements, which are accounting for 80% of the nodes.

Who should be my Facebook partner? ...

Szolnoki, Gulyás and Kampis

4.2 The giant component

The giant component is 1.5 times denser than is the whole network, its density is 0.06. Its diameter is still 18, but here the average path length is 11.5, which is almost two times larger than that of the whole network, which is due to the fact that some nodes with large a betweenness centrality values are not part of the giant component.

4.2.1 Clusters

We have used the "modularity" command in Gephi to identify clusters within the giant component. The calculation of the index is based on the comparison of the network's data and the data of an equivalent network where the edges are randomly generated. By this, the modularity index detects clusters with looking for "unexpectedly" few connecting edges between clusters and "unexpectedly" many nodes within clusters [3].

Gephi uses the algorithm introduced in Blondel et al.'s paper [7]. It asks for a resolution (r) value based on the article by Lambiotte et al. [6]. The higher the value of r , the less number of clusters are formed by the modularity command. Furthermore this algorithm is not deterministic, it may provide different results each time it is run. Therefore, it is reasonable to run the command for many different r values and also it is important to run the command many times with the same resolution value as well.

For $r=1$, the modularity command provided results with cluster numbers varying between 5 and 6. In these cases, there remained some large clusters with very mixed elements, therefore we decided to lower the r value and this way to increase the number of clusters in the results. In case of $r=0.5$, for instance, the results provided 8-10 clusters, which we considered as unnecessarily too many, as in these cases even well-formed clusters were further divided. Finally we ended up with a result where 7 clusters were formed at $r=0.75$.

During the scenario analyses (in the case of the 5-6-7 cluster results) there was one very stable cluster which contained the same elements in each scenario. In this cluster there are 17 elements and the Facebook pages are mostly related to gastro sites and baby-mother magazines. These pages are complemented by some strongly related brands' sites and a tabloid-science magazine that covers related topics. These pages have a well defined target group: women and foremostly mothers.

The second most stable cluster which has only slightly changed during the scenarios is contained of 12-13 FB pages of video games - soccer magazines - soccer clubs - sports betting - quotes and sayings - and the largest Hungarian jokes page. This cluster contains pages targeted mostly towards men and have again well focused topics: games, sports, betting and the accompanying quotes and jokes.

The third, still fairly stable group with 16-18 elements contains FB pages of public figures from the music industry and the tabloids. These pages are accompanied by pages of related brands (radio channels, TV series, fitness). This cluster is also a well focused group: namely, media.

The size of the fourth cluster frequently varied between 18 and 26 FB pages depending on the analysis details, but its core was stable: news sites and political parties. These in some cases were complemented by FB pages of fake-science-fantastic news sites. Surprisingly this cluster contained parties and politicians from all political views.

The fifth-sixth-seventh clusters were less stable. These all contain pages from the topics of entertainment and leisure: women's magazines, tabloid news sites, leisure related sites and fake-science-fantastic sites. These Facebook pages formed different clusters in the different scenarios. The result we chose forms three clusters. There is one cluster containing 7 Facebook pages from

Who should be my Facebook partner? ...

Szolnoki, Gulyás and Kampis

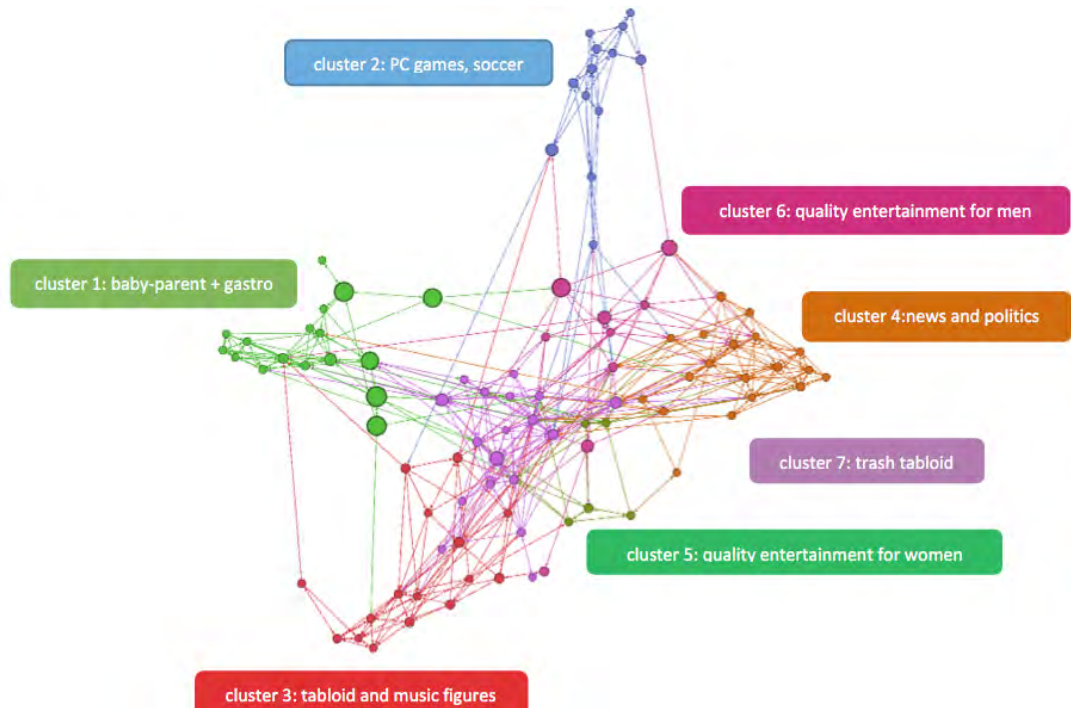


Figure 3: The final clusters.

quality women magazines, classical music, traveling and male singers, these entertainment/relaxation pages target women who are more educated/wealthier. There is another cluster with 7 Facebook pages that are very similar to the previous groups' but target male users: men's magazines, radio channels, animal lovers. And finally, there remained an 18-element group which contains trash tabloid products: tabloid news sites, women tabloid magazines, female singers, showmen, and fake-fantastic news sites.

5 Discussion

From the analysis of the strong components and mostly of the giant component, it seems obvious that the topic is the primary organizing principle. Surprisingly topic is even stronger than political views, i.e. those who are interested in politics are more active on the pages of political or news sites from different political views than on other sites like gastro, tabloid, etc. Another interesting result is that ownership structure does not affect the relationship between pages. This means that it is hard for the operators to direct users to their other sites if those sites are not related in topic.

To the research question, what are the organizing principles within the given topics, the cases of the 5th 6th and 7th clusters provide an answer. In all these three clusters, the topic is entertainment and leisure. The clusters differ in the positioning of this topic within the gender of the users and social status. These two could be sub-organizing principles therefore within a topic.

A further interesting result is that while according to the MTE study [1] celebrities are not

Who should be my Facebook partner? ...

Szolnoki, Gulyás and Kampis

a relevant topic for users, in our analysis they formed a separate cluster. Another interesting result was that while in the MTE study all user types indicated funny Facebook pages as their most favorite site, these sites do not form a separate cluster in themselves, rather they join other pages in different clusters that are determined by topics. The same happens in the case of quotes sites.

Regarding related topics for identifying link-exchange partners, the clusters provide straightforward suggestions for Facebook page operators. For Baby-mother sites for example, it could be beneficial to partner with gastro sites. For sports betting sites PC game sites could be good partners, and so on.

6 Summary

In this article we aimed to analyze how the large-scale Hungarian Facebook pages relate to each other, what is their primary ordering principle. We have built a network database, where the nodes are Facebook pages and the edges represent the relation that a given Facebook page has with other Facebook pages when they share a common active user fan base. This resulted in a directed network with 121 nodes and 655 edges. It turned out that even in the case of Facebook pages with a large fan base there are few with many connections and many with few connections, i.e. the indegree distribution is close to power distribution.

The network was weakly connected, with 14 components. The giant component consisting of 80% of the nodes could be divided into stabile clusters. After the scenario analysis we chose a result with 7 clusters.

Among the clusters the most stable were the 1. baby-mother-gastro topic cluster, the 2. sports betting - PC games - jokes site, the 3. tabloid and music public figures, and the 4. news and politics clusters. The remaining pages were all in the topic of entertainment and leisure. Within this large group my 7 cluster result identified three different clusters, one for high-status women, one for high-status men and one where trash tabloid topic dominated.

Based on the results, the primary ordering principle is the topic, while ordering principles such as ownership background, political views could not be identified. Within large topics gender and social background could be identified as sub ordering principles.

The clusters we have identified could provide useful input for Facebook page operators in finding valuable link-exchange partners.

7 Acknowledgement

This paper is partially supported by the Russian Scientific Foundation, proposal #14-21-00137, "Supercomputer simulation of critical phenomena of complex social systems".

Who should be my Facebook partner? ...

Szolnoki, Gulyás and Kampis

References

- [1] Facebook hasznalati szokások. kutatási jelentés az mte számára 2015. november, 2016. http://mte.hu/wp-content/uploads/2016/02/MTE_FB_201511_v2.pptx.
- [2] FPK. Engagement. <https://www.fanpagekarma.com/help#?Engagement>.
- [3] Newman M. E. J. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, 2006. <http://www.pnas.org/content/103/23/8577.full.pdf>.
- [4] Ingram M. Facebook has taken over from google as a traffic source for news. *Fortune.com*, 2015. <http://fortune.com/2015/08/18/facebook-google/>.
- [5] Bilton R. In search of facebook love, publishers form link-sharing pacts with each other. *Digiday*, February 4, 2016. <http://digiday.com/publishers/social-world-publishers-taking-link-sharing-facebook/>.
- [6] Lambiotte R., Delvenne J.-C., and Barahona M. Laplacian dynamics and multiscale modular structure in networks *ieee transactions on network science and engineering. IEEE Transactions on Network Science and Engineering*, 1(2):76–90, 2015.
- [7] Blondel D. V., Guillaume J.-L., Lambiotte R., and Lefebvre E. Fast unfolding of communities in large networks. *J. Stat. Mech.*, page P10008, 2008. <http://digiday.com/publishers/social-world-publishers-taking-link-sharing-facebook/>.

EXHIBIT 14

Original Paper

Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis

Amit Baumel¹, PhD; Frederick Muench², PhD; Stav Edan¹; John M Kane³, MD

¹Department of Community Mental Health, University of Haifa, Haifa, Israel

²The Partnership for Drug-Free Kids, New York, NY, United States

³Psychiatry Research, The Zucker Hillside Hospital, Northwell Health, Glen Oaks, NY, United States

Corresponding Author:

Amit Baumel, PhD

Department of Community Mental Health

University of Haifa

Abba Khoushy Ave 199

Haifa, 3498838

Israel

Phone: 972 048240111

Email: abaumel@univ.haifa.ac.il

Abstract

Background: Understanding patterns of real-world usage of mental health apps is key to maximizing their potential to increase public self-management of care. Although developer-led studies have published results on the use of mental health apps in real-world settings, no study yet has systematically examined usage patterns of a large sample of mental health apps relying on independently collected data.

Objective: Our aim is to present real-world objective data on user engagement with popular mental health apps.

Methods: A systematic engine search was conducted using Google Play to identify Android apps with 10,000 installs or more targeting anxiety, depression, or emotional well-being. Coding of apps included primary incorporated techniques and mental health focus. Behavioral data on real-world usage were obtained from a panel that provides aggregated nonpersonal information on user engagement with mobile apps.

Results: In total, 93 apps met the inclusion criteria (installs: median 100,000, IQR 90,000). The median percentage of daily active users (open rate) was 4.0% (IQR 4.7%) with a difference between trackers (median 6.3%, IQR 10.2%) and peer-support apps (median 17.0%) versus breathing exercise apps (median 1.6%, IQR 1.6%; all $z \geq 3.42$, all $P < .001$). Among active users, daily minutes of use were significantly higher for mindfulness/meditation (median 21.47, IQR 15.00) and peer support (median 35.08, $n=2$) apps than for apps incorporating other techniques (tracker, breathing exercise, psychoeducation: medians range 3.53-8.32; all $z \geq 2.11$, all $P < .05$). The medians of app 15-day and 30-day retention rates were 3.9% (IQR 10.3%) and 3.3% (IQR 6.2%), respectively. On day 30, peer support (median 8.9%, $n=2$), mindfulness/meditation (median 4.7%, IQR 6.2%), and tracker apps (median 6.1%, IQR 20.4%) had significantly higher retention rates than breathing exercise apps (median 0.0%, IQR 0.0%; all $z \geq 2.18$, all $P \leq .04$). The pattern of daily use presented a descriptive peak toward the evening for apps incorporating most techniques (tracker, psychoeducation, and peer support) except mindfulness/meditation, which exhibited two peaks (morning and night).

Conclusions: Although the number of app installs and daily active minutes of use may seem high, only a small portion of users actually used the apps for a long period of time. More studies using different datasets are needed to understand this phenomenon and the ways in which users self-manage their condition in real-world settings.

(*J Med Internet Res* 2019;21(9):e14567) doi: [10.2196/14567](https://doi.org/10.2196/14567)

KEYWORDS

user engagement; usage; adherence; retention; mental health; depression; anxiety; mHealth

Introduction

The wide dissemination of mobile phone devices and the leap in the development and distribution of mobile health (mHealth) apps have altered the ways in which scholars conceptualize care management in the behavioral health domain. The conversation has shifted from patients and providers to individuals who can now engage in self-care around the clock outside of traditional health care settings (eg, [1-3]). Approximately 77% of the US adult population, and more than 89% of those younger than 50 years, now own a mobile phone [4,5] where they can store and use computerized apps. This widespread use has established a market for mHealth apps. Accordingly, a 2015 World Health Organization survey identified approximately 15,000 mobile apps for health care, with at least 29% designed for mental health [6].

The use of unguided apps has the potential to increase access to care in a scalable manner by reducing the costs associated with service uptake [7,8]. However, the impact of digital interventions is limited by their ability to engage users in therapeutic activities and to support user adherence to the therapeutic process [9,10]. Digital interventions require individuals to engage with self-care outside of traditional settings; therefore, individuals' engagement must compete with other events in their daily lives and endure fluctuating motivation to be involved in effortful behavior [11]. As a result, user engagement with mobile apps and websites across the behavior change spectrum is low in the absence of human support [12-14]. Furthermore, various studies have suggested that most users of unguided Web-based programs exit websites before the full completion of the offered program [9,10,15,16]. For example, Christensen and colleagues [17] reported that less than 1% of users completed all modules in MoodGym, an open-access website for depression. In a systematic review of published articles reporting real-world user engagement with unguided programs for depression, anxiety, or mood enhancement, Fleming and colleagues [18] reported that 7% to 42% of users of Web- and app-based programs engaged in moderate use (completing between 40% and 60% of modular fixed-length programs or continuing to use the app after 4 weeks). For example, the developers of the PTSD Coach mobile app reported a usage decline over time, with 41.6% continuing to use the app 1 month after installation and 19.4% after 6 months [19]. Among Happify mobile app users, 3.5% completed a 6-week assessment. However, the authors noted that these users might have completed assessments without engaging in other content [20] (see [18] for a review).

Understanding patterns of real-world usage of e-mental health apps outside of empirical trials is key to maximizing the potential of apps to increase the public self-management of care. Utilization in real-world settings may differ from that in study settings for several reasons. First, empirical study settings include enrollment and assessment procedures that are not part of real-world utilization of the app, as trials largely emphasize internal validity over real-world generalizability [13]. Ebert and Baumeister [21] claim, for example, that within randomized trials "the securing of commitment represents an adherence-promoting element in self-help interventions." It is

reasonable to assume that the human contact provided by research coordinators, provision of ongoing assessments, and reimbursement to incentivize the completion of assessments—none of which are available in real-world use—impact engagement patterns with the interventions. Second, from an external validity perspective, recruitment challenges in trials are often addressed by increasing the reach to potential participants through the expansion of participating venues and the refinement of social media strategies [13]. In this way, researchers unintentionally recruit people who are much more likely to adhere to e-mental health technologies than people in the general population who download and try available programs "in the wild." Such assumptions are supported by a systematic review of internet interventions for anxiety and depression, which found that the rates of attrition in randomized controlled trials were lower than the reported dropout rates from open-access websites [22].

Overall, there is a need to understand how the general population engages with the most popular unguided mobile apps targeting anxiety, depression, or emotional well-being, and whether there is a difference in how individuals engage with these apps depending on the mental health focus or incorporated techniques. Although some developer-led studies have published results on the use of individual mental health apps deployed in real-world settings, to the best of our knowledge, no study has examined a large sample of mental health apps relying on independently collected data. This investigation is feasible by leveraging the big data commonly generated and stored by digital platforms that record user traffic in the wild [23,24]. Leveraging such data, this examination provides benchmarks of app usage in the real world, where the general public is expected to benefit from their engagement with unguided programs. This information could shed light on specific engagement problems and opportunities for new intervention development and may offer a resource for researchers and developers who want to study and compare their app performance with similar apps.

For this study, a panel provided objective aggregated nonpersonal data on user engagement with mobile apps to analyze patterns of mental health app usage. The three primary aims were to (1) describe common usage patterns of popular unguided apps based on available metrics, (2) identify patterns of user retention over the first 30 days after app installation, and (3) explore whether these patterns differ based on the app's mental health focus and primary incorporated techniques.

Methods

Search Strategy

The search strategy aimed at identifying the most-installed unguided apps targeting depression, anxiety-related problems, or mental health. We used keywords related to depression and anxiety because of the high prevalence of these conditions [25,26]. We also included mental health apps that focused on happiness or the enhancement of mental health (ie, mindfulness meditations) because our previous work identified them as highly popular mental health tools [27,28]. We conducted a systematic engine search of the Google Play Store in November 2018 using the following terms: "depression" OR "mood" OR

“anxiety” OR “panic attack” OR “phobia” OR “social phobia” OR “PTSD” OR “posttraumatic stress disorder” OR “stress reduction” OR “worry relief” OR “OCD” OR “obsessive compulsive disorder” OR “mental health” OR “emotional well-being” OR “happiness.” One researcher documented all the apps emerging from the first 100 search results of each keyword, removed duplicates, and sorted them alphabetically. We also included a manual search of apps presented on MindTools.io [27] and PsyberGuide [29].

App Screening and Inclusion Criteria

Determining Apps' Number of Installs Threshold

To avoid including apps without a representative number of users, and to determine a minimum threshold for inclusion, we assessed the install categories presented by Google Play based on the number of app installs (eg, 10,000, 50,000 installs). Table 1 presents a preliminary analysis of the number of identified

apps in each install category and the aggregated minimum number of app installs and corresponding percentages. Included apps had at least 5000 installs after removing any nonrelevant apps based on their title (ie, apps that were clearly not targeted at emotional well-being such as Heart Rate Monitor & Pulse Checker, 7 Minute Workout, 30 Day Fitness Challenge). Adding all the apps in the 5000 installs category would have resulted in a less than 0.5% increase in the total sample of users. Therefore, we determined an inclusion threshold of 10,000 app installs. Table 1 also shows that a small number of apps within the higher install categories were responsible for the most app installs. To make sure that including a large portion of apps with a relatively smaller number of installs (eg, <10,000 app installs) would not bias the results, we also examined whether there was a difference in the pattern of results based on the number of app installs. This will be further explained in the data analysis section.

Table 1. Analysis of install categories based on the number of apps in each category.

| Install category | Apps identified, n | Minimum identified app installs within this category ^a , n | Cumulative frequency of app installs based on category threshold ^b , n | Added percentage of installs to the overall sample ^c , % |
|---------------------|--------------------|---|---|---|
| ≥10,000,000 | 2 | 20,000,000 | 20,000,000 | 100.00 |
| 5,000,000-9,999,999 | 6 | 30,000,000 | 50,000,000 | 60.00 |
| 1,000,000-4,999,999 | 21 | 21,000,000 | 71,000,000 | 29.58 |
| 500,000-999,999 | 23 | 11,500,000 | 82,500,000 | 13.94 |
| 100,000-499,999 | 69 | 6,900,000 | 89,400,000 | 7.72 |
| 50,000-99,999 | 33 | 1,650,000 | 91,050,000 | 1.81 |
| 10,000-49,999 | 103 | 1,030,000 | 92,080,000 | 1.12 |
| 5000-9999 | 66 | 330,000 | 92,410,000 | 0.36 |

^aThe number of apps multiplied by the minimum number of installs based on the install category.

^bThe accumulated number of app installs in all install categories above and including the current install category.

^cThe added percentage of installs to the total sample if the current install category is added to the analysis; it represents the percentage of the total number of app installs within this category divided by the accumulated number of app installs based on the current category threshold.

Inclusion and Exclusion Criteria

To be included in this review, apps had to:

1. Be in English;
2. Have at least 10,000 installs documented on Google Play;
3. Focus on mental illness, mental health, or emotional well-being not specifically related to another medical condition (for example, we excluded apps specifically focused on stress reduction due to a physical medical issue such as heart attack); and
4. Incorporate recognized techniques aimed at promoting self-management of mental health problems such as coping with negative symptoms (eg, feeling nervous, loss of energy), achieving positive results (eg, feeling better), or symptom management (eg, mood tracking). We excluded apps focused on the incorporation of sham techniques (see Multimedia Appendix 1 for a definition of sham techniques).

We excluded apps that:

1. Required payment for installation or provided a free trial only for a limited amount of time because it would be expected to bias program usage (free to install apps that included in-app purchases were not excluded);
2. Were therapist-based (eg, telepsychiatry) because the study was focused on unguided interventions; and
3. Were not meant to be used for more than a few times (eg, tests, one-time exposure technique) or were merely magazines.

Two independent reviewers screened the apps based on the inclusion and exclusion criteria. All disagreements were discussed with a third author with reference to the apps until consensus was reached.

Coding

Two independent reviewers coded the apps' incorporated techniques based on the following categories: mindfulness/meditation, tracker (including diary or journal), psychoeducation, peer support, and breathing exercise (not exercised as part of a meditation program). These categories were based on previous work done on the therapeutic

components of mental health apps [27,30], drawing on the thematic analysis method suggested by Braun and Clarke [31]. The categories were designed to represent nonoverlapping components of potential therapeutic engagement (see [Multimedia Appendix 2](#) for definitions of categories). Although our goal was to identify how specific techniques related to patterns of app use, our metrics did not enable us to differentiate between various techniques incorporated within the same app (ie, we could not tell which parts in the app the users were using). Therefore, we also added a coding of “primary technique” in cases where the app mostly incorporated one technique that was deemed to be the main reason for the app’s use (eg, mindfulness/meditation). It is important to note that this limitation did not enable us to include app features that might influence user engagement but were not identified as a primary incorporated technique. Similarly, it was not feasible to target specific theoretical modalities, such as cognitive behavioral therapy. Because nearly all apps included some components of cognitive behavioral therapy, these were impossible to dismantle given our data.

An app’s mental health focus was determined in the following manner: first, the app’s description had to explicitly state that it targeted people with [mental health focus] and, second, most of the techniques used within the app had to have been built to help users cope with or manage their symptoms directly related to the mental health focus. We grouped apps based on several mental health foci. Under “mental health problems,” we included apps that were focused on supporting people coping with depression, anxiety-related disorders, and emotional difficulties. We also subcoded the app with the terms (a) anxiety-related disorders or (b) depression if the app specifically targeted only one of these aims. (During our coding process, we did not identify another theme for the remaining apps.) Under “happiness,” we included apps that focused on nurturing happiness or general positivity (eg, exercising gratitude, happiness assessment, suggestions for activities nurturing positive feelings), rather than the management of mental health states or problems.

During our coding process, we found a greater ambiguity around the description of apps with a primary incorporated technique of mindfulness/meditation, which leaned more toward enhancing emotional well-being (ie, helping users achieve a positive sense of experience and good mental health), but also aimed at stress reduction. Therefore, we grouped mindfulness and meditation apps separately and did not attribute either of the two mental health foci to them. For this reason, and to enable a proper comparison between categories, we present the mindfulness/meditation category in both the mental health focus and technique outcomes, despite being the same results.

A Cohen kappa interrater agreement of .92 was obtained for coding the variables of interest (incorporated technique, primary technique, and mental health focus). All disagreements were discussed with a third author with reference to the apps until consensus was reached.

Behavioral Data on User Engagement in the Real World

Information on user traffic was obtained from SimilarWeb’s Pro panel data [32]. The panel provides aggregated nonpersonal information on user engagement with websites and mobile apps all over the world to enable Web and mobile app traffic research and analytics. The panel is based on several sources of anonymized usage data, such as data obtained from consenting users of mobile apps (ie, products). A dedicated product team at SimilarWeb is responsible for building and partnering with hundreds of high-value consumer products that make up the panel. According to SimilarWeb, the products are used across diverse audiences, without cluttering the user with advertisements. While benefiting from the products, users contribute to the panel because they enable the documentation of their online or mobile app usage activities seamlessly and anonymously [32]. The data are not used by SimilarWeb or provided to any third parties for the purposes of marketing, advertising, or targeting of individual subjects. The data-gathering procedures comply with data privacy laws, including the way data are collected, anonymized, stored, secured, and used. These procedures are updated regularly based on evolving data privacy legislation and requirements, such as the European Union’s General Data Protection Regulation [33].

Our examination of data validity was tested and presented in a previous study [28]. An Oath researcher [34] (RW) examined 30 randomly selected mobile apps with data on SimilarWeb and usage data in Oath’s independent records. The researcher examined the correlation between the average number of user sessions per day in the two datasets, finding a very strong Spearman correlation ($N=30$, $r=.77$, $P<.001$). In our study, we also examined the Spearman correlation between app install categories presented on Google Play (eg, 10,000, 50,000) and the number of downloads documented on SimilarWeb, and found a very strong correlation ($N=93$, $r=.81$, $P<.001$). These findings suggest a sufficient convergent validity, which is recommended to be above .70 [35].

The study was approved by University of Haifa Institutional Review Board, Haifa, Israel. The measures were set to include data gathered over a 12-month period from August 1, 2017, to July 31, 2018. For each app, available metrics on the panel included app open rate (the average percentage of daily active users out of the total sample of people who currently had the app installed), average number of sessions in a day per daily active user, and average daily minutes of use per daily active user. User 30-day retention included the percentage of users who opened the app each day between day 1 and day 30 out of the number of users who installed and opened the app on day 0. Usage patterns by time were available only for apps with a very large number of users. It was represented by two metrics—average percentage of use per hour (24 hours; eg, 7:00 am, 8:00 am) and per day (7 days; eg, Sunday, Monday)—both calculated based on total app usage.

Data Analysis

We did not assume a normal distribution of the metrics; therefore, medians and interquartile ranges (IQRs) were used as descriptive statistic measures. In cases in which a category

included a small number of apps ($n \leq 5$), we used range instead of IQR. To examine differences in usage metrics between apps with different mental health foci or techniques, a Kruskal-Wallis one-way analysis of variance (ANOVA) was performed, followed by Mann-Whitney U tests to identify the source of the difference. To examine dependencies in the distribution of categorical values in relevant cases, we used chi-square tests. Most app installs came from a small number of apps with a large number of installs (see Table 1), so we conducted a sensitivity analysis to examine whether including apps with a smaller number of installs would bias the results. Mann-Whitney U tests were conducted to compare the distributions of the usage patterns for the top 5 installed apps and the remaining apps from each category presented in the results section (and that included more than five apps). We picked the top 5 apps based on their install category in Google Play. In cases in which several apps “competed” for the fifth place in the same install category, the app with the higher number of downloads (as documented in the SimilarWeb user panel) was chosen.

Results

Screening

Figure 1 presents the app inclusion flow diagram. The engine search and manual searches produced a total of 386 apps with

10,000 installs or more. Through the first screening process, 299 apps were identified and accessed for a detailed evaluation, and 93 apps were finally included in this study analysis (see Multimedia Appendix 3 for a complete list of included apps).

Description of Apps

The mental health focus of 59 (63%) apps was a mental health problem. Of these, 19 focused specifically on anxiety-related disorders and 4 focused specifically on depression. In addition, 8 (9%) apps focused on happiness, and 26 (28%) apps focused on the enhancement of emotional well-being through mindfulness/meditation. The distribution of apps based on incorporated techniques is presented in Table 2. Overall, 60 of 93 (65%) apps had a primary incorporated technique, and 33 (36%) apps had two or more incorporated techniques, none of which were primary. Mindfulness/meditation was the most frequent technique as the primary technique of the app (26/93, 28%), followed by use of a tracker (22/93, 24%). Psychoeducation (35/93, 38%) was the most frequent salient technique to be used not as the primary technique, followed by use of a tracker (28/93, 30%).

Figure 1. App inclusion flow diagram.

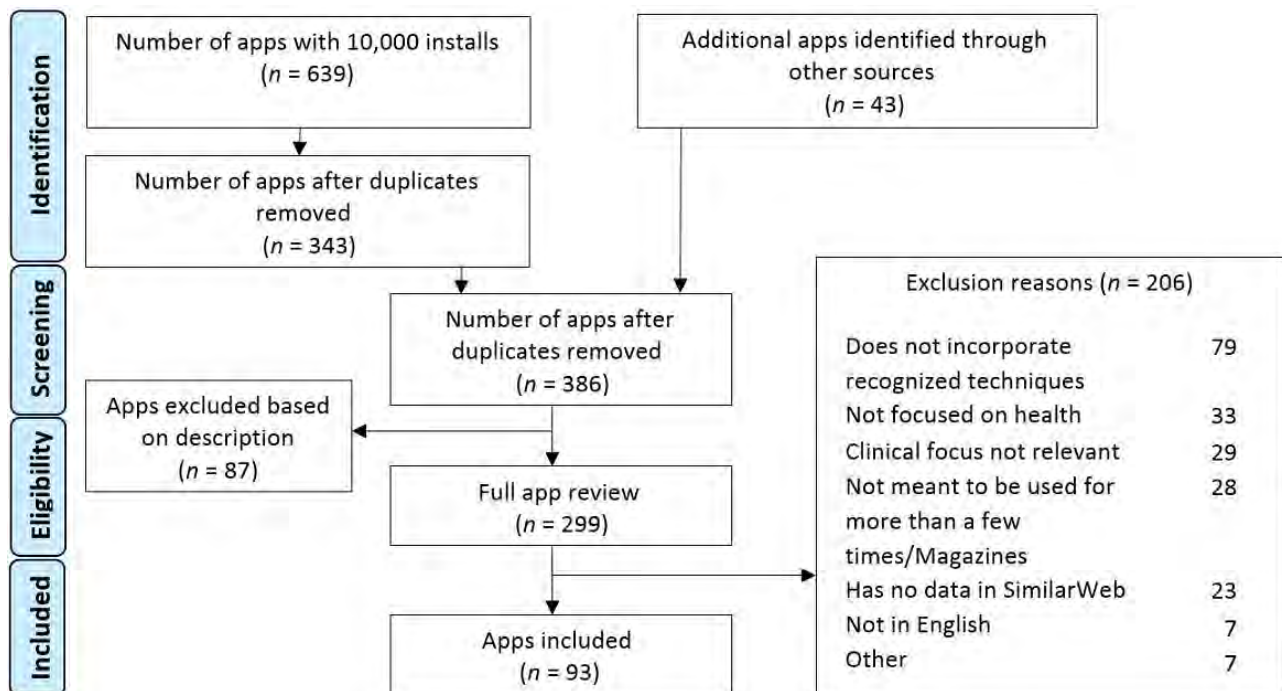


Table 2. Distribution of incorporated techniques in the app sample (N=93).

| Incorporated technique | Primary technique, n (%) | Cotechnique ^a , n (%) | Total, n (%) |
|------------------------|--------------------------|----------------------------------|--------------|
| Mindfulness/meditation | 26 (28) | 14 (15) | 40 (43) |
| Tracker | 22 (24) | 28 (30) | 50 (54) |
| Breathing exercise | 7 (8) | 20 (22) | 27 (29) |
| Psychoeducation | 3 (3) | 35 (38) | 38 (41) |
| Peer support | 2 (2) | 7 (8) | 9 (1) |

^aThe technique is saliently presented in the app but is not considered a primary technique.

App Usage by Daily Active Users

All apps had complete metrics on app usage by daily active users. Medians and IQRs of daily app usage are presented in Table 3 based on the app's mental health focus and in Table 4 based on the app's incorporated techniques. As shown in Table 3, the median app open rate was 4.0% (IQR 4.7%), with medians of 3.28 (IQR 2.53) daily sessions and 13.03 (IQR 14.27) minutes of app use per active user. Daily active usage of mindfulness/meditation apps (median 21.47, IQR 15.00) was found to be significantly higher than the usage of apps for mental health problems (median 10.02, IQR 10.60; $z=4.64$, $P<.001$) or for happiness (median 7.77, IQR 6.90; $z=3.82$, $P<.001$). No

other significant difference in app usage was found between mental health foci, including between anxiety- and depression-related apps. As seen in Table 4, the number of app minutes of use was significantly higher for mindfulness/meditation (median 21.47, IQR 15.00) and peer support (median 35.08, $n=2$) than for other techniques (all $z \geq 2.11$, all $P<.05$). In addition, tracker (median 6.3%, IQR 10.2%) and peer support (median 17.0%, $n=2$) apps had significantly higher open rates than breathing exercise apps (median 1.6%, IQR 1.6%; all $z \geq 3.42$, all $P<.001$). No significant differences in usage patterns were found for apps without a primary strategy that incorporated more than one technique.

Table 3. App usage based on app mental health focus (N=93).

| Mental health focus | Apps, n | Installation category, median (IQR) | Open rate (%), median (IQR) | Daily number of sessions per active users, median (IQR) | Daily minutes of use per active user, median (IQR) ^a |
|-------------------------------------|---------|--|-----------------------------|---|---|
| All apps | 93 | 100,000 (90,000) | 4.0 (4.7) | 3.28 (2.53) | 13.03 (14.27) |
| Mental health problems | 59 | 50,000 (90,000) | 4.0 (5.1) | 3.77 (3.15) | 10.02 (10.60)* |
| Anxiety | 19 | 10,000 (40,000) | 2.6 (2.5) | 3.58 (3.49) | 08.17 (09.42) |
| Depression | 4 | 100,000 (50,000-100,000 ^b) | 4.8 (3.0-6.8 ^b) | 5.22 (3.97-6.55 ^b) | 06.97 (02.05-15.12 ^b) |
| Happiness | 8 | 100,000 (50,000) | 3.7 (5.3) | 3.50 (4.18) | 7.77 (6.90)* |
| Mindfulness/meditation ^c | 26 | 100,000 (650,000) | 4.1 (3.3) | 2.96 (1.66) | 21.47 (15.00)** |

^aCategories with different number of asterisks (*, **) within a column are significantly different ($P<.05$) based on our analytical approach, which included Kruskal-Wallis one-way ANOVA at the variable level, followed by Mann-Whitney U tests.

^bDue to a small number of included apps, brackets in this cell reflect the range (minimum-maximum value) and not the IQR.

^cMindfulness/meditation is presented as a separate mental health focus because all apps in this category were not attributed to another focus as they focus on enhancement of well-being as well as stress reduction.

Table 4. App usage based on app incorporated technique (N=93).

| Incorporated technique | Apps, n | Installation category, median (IQR) | Open rate (%), median (IQR) ^a | Sessions per active user, median (IQR) | Daily minutes of use per active user, median (IQR) ^a |
|-------------------------------------|---------|---------------------------------------|--|--|---|
| Primary technique | | | | | |
| Mindfulness/meditation | 26 | 100,000 (650,000) | 4.1 (3.3) | 2.96 (1.66) | 21.47 (15.00)* |
| Tracker | 22 | 50,000 (90,000) | 6.3 (10.2)* | 4.58 (4.47) | 07.27 (08.83)** |
| Breathing exercise ^b | 7 | 10,000 (40,000) | 1.6 (1.6)** | 2.19 (1.23) | 08.32 (19.02)** |
| Psychoeducation | 3 | 10,000 (10,000-100,000 ^b) | 3.0 (2.5-3.3 ^c) | 4.16 (2.57-4.80 ^c) | 03.53 (02.07-19.23 ^c)** |
| Peer support ^d | 2 | 300,000 (N/A ^e) | 17.0 (N/A)* | 8.67 (N/A) | 35.08 (N/A)* |
| Number of primary techniques | | | | | |
| 2 techniques | 17 | 50,000 (90,000) | 4.0 (5.6%) | 3.18 (1.40) | 07.83 (11.93) |
| ≥3 techniques ^f | 16 | 100,000 (50,000) | 3.2 (3.1%) | 4.06 (3.91) | 12.88 (07.13) |

^aCategories with different number of asterisks (*, **) within a column are significantly different ($P < .05$) based on our analytical approach, which included Kruskal-Wallis one-way ANOVA at the variable level, followed by Mann-Whitney U tests.

^bNot including mindfulness/meditation.

^cDue to the small number of included apps, brackets in this cell reflect the range (minimum-maximum value) and not the IQR.

^dDue to the small number of included apps, IQR or range could not be calculated (marked with N/A).

^eN/A: not applicable.

^fIncludes two apps that use a chatbot (Wysa, Woebot), which did not have a different pattern of results emerging for a certain direction.

User 30-Day Retention

Fifty-nine apps (63%) had data on user retention. Chi-square tests for independence revealed no difference between apps with or without user retention data in the distribution of mental health foci ($\chi^2_2=2.1$, $P=.36$) and primary incorporated techniques ($\chi^2_4=3.8$, $P=.44$). Figure 2 presents user 30-day retention by the app's mental health focus; Figure 3 presents user 30-day retention by the app's incorporated technique. In both figures, there is a sharp decline of more than 80% in app open rates between day 1 and day 10, whereas the differences between day 15 and day 30 are smaller and represent a decline of approximately 20% in app open rates. Figure 2 reveals that, relative to users who opened the app on day 0, the median app open rate was as follows: 69.4% (IQR 27.8%) of users opened

it on day 1, 3.9% (IQR 10.3%) of users opened it on day 15, and 3.3% (IQR 6.2%) of users opened it on day 30. Kruskal-Wallis one-way ANOVAs revealed no significant differences in app open rates on day 30 based on mental health focus ($H_2=1.88$, $P=.39$) and a significant difference in app open rates on day 30 based on incorporated technique ($H_5=11.31$, $P=.046$). Mann-Whitney U tests revealed that on day 30 peer support (median 8.9%), mindfulness/meditation (median 4.7%, IQR 6.2%), and tracker/diary apps (median 6.1%, IQR 20.4%) had significantly higher retention rates than breathing exercise apps (median 0.0%, IQR 0.0%; all $z \geq 2.18$, all $P \leq .04$). This pattern of difference is also descriptively apparent in 15-day retention, in which the median retention for breathing exercise apps was 0.0% (IQR 0.0%), whereas the range of medians for peer support, mindfulness/meditation, and tracker/diary apps was from 4.9% (IQR 7.1%) to 11.9% (IQR 0.7%).

Figure 2. App 30-day retention by mental health focus. The percentages reflect the number of users who opened the app from day 1 to day 30 out of the number of users who installed and opened the app on day 0.

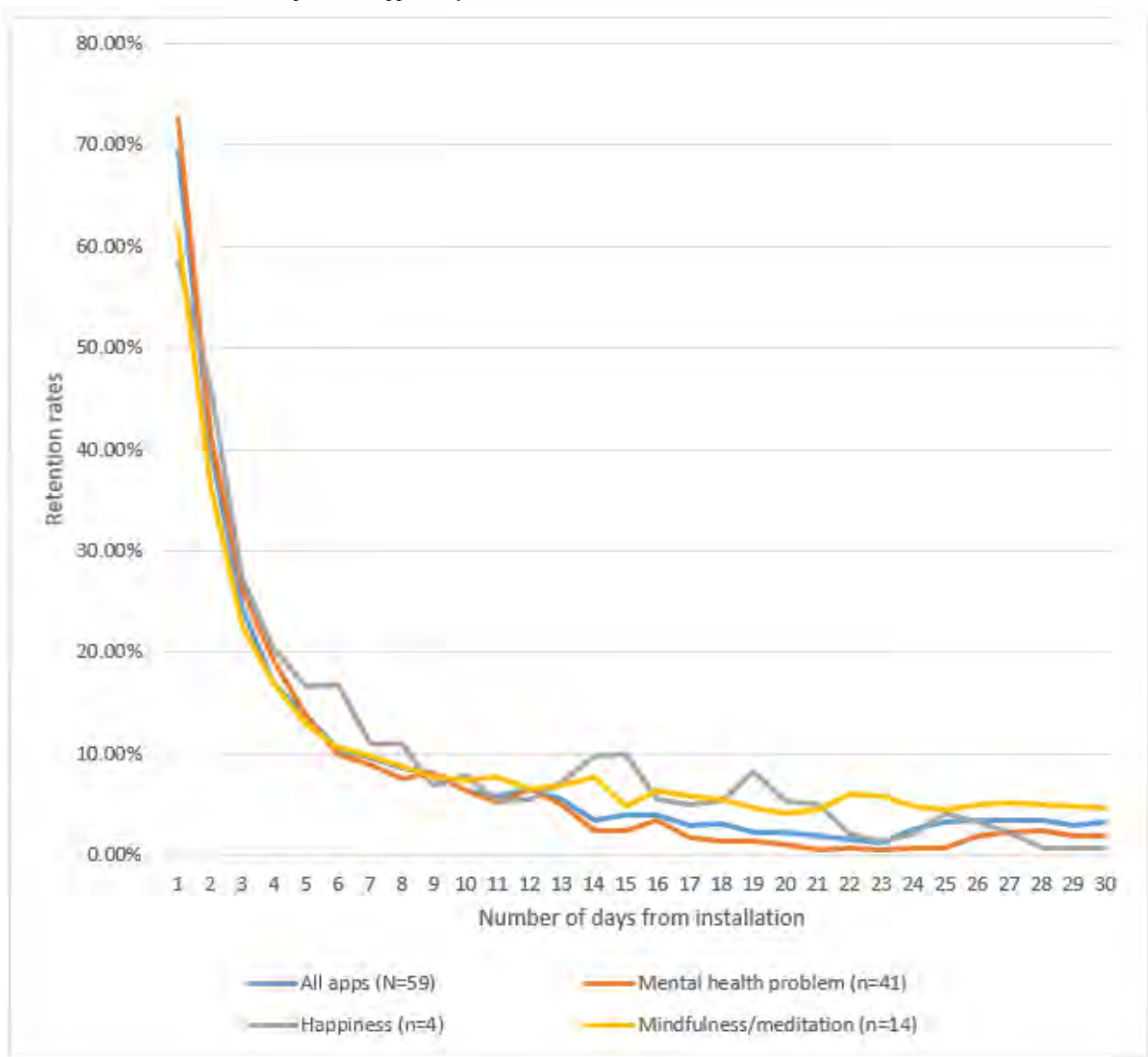
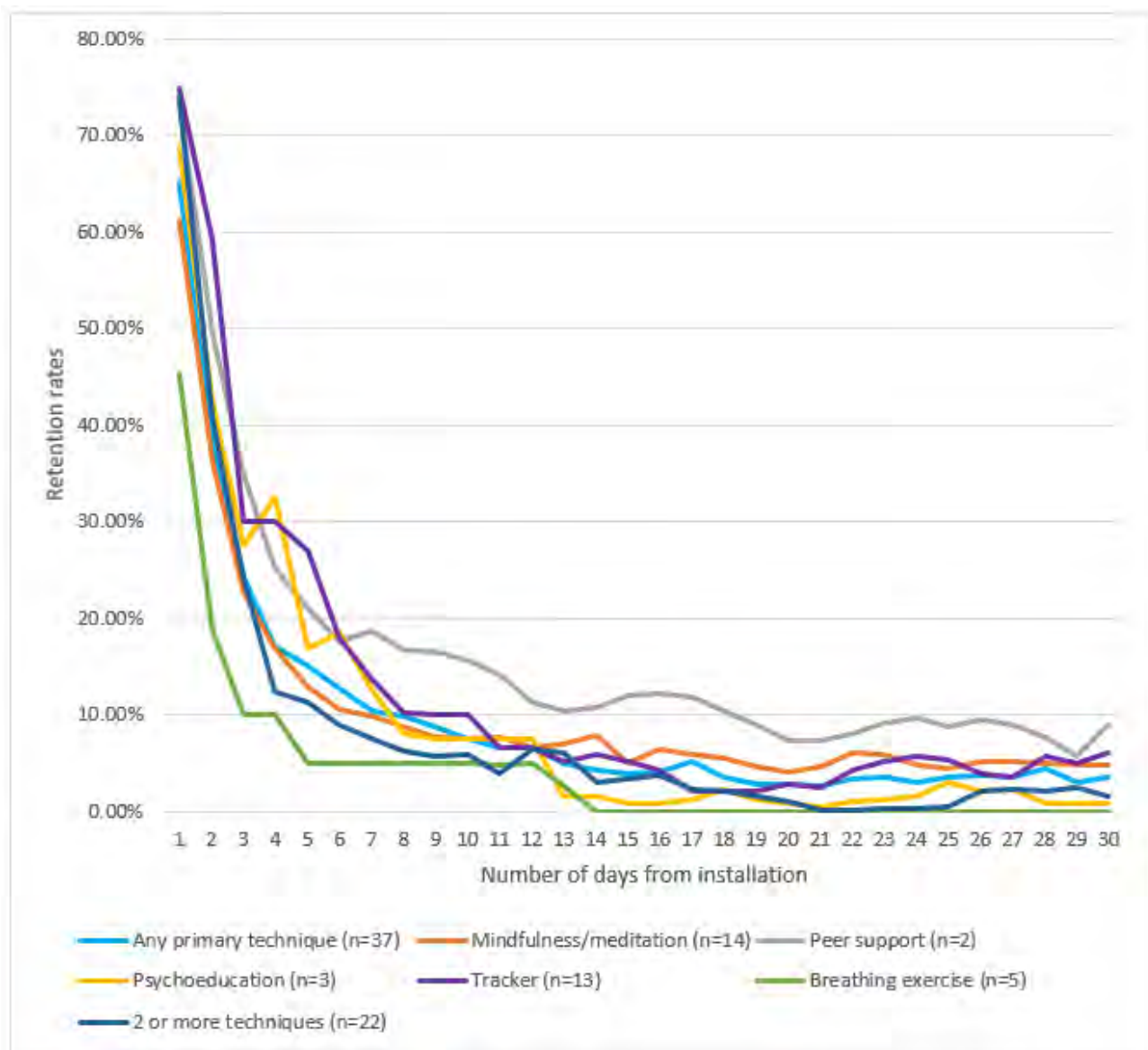


Figure 3. App 30-day retention by primary incorporated technique. The percentages reflect the number of users who opened the app from day 1 to day 30 out of the number of users who installed and opened the app on day 0.



Usage Pattern by Hours and Days

Sixteen apps had data on hourly and daily app usage. Figure 4 presents the hourly usage patterns of apps and Figure 5 presents the daily usage patterns of apps. The number of apps with available data was small; therefore, we only present categories with data on more than three apps. Furthermore, we have not conducted statistical testing to compare program usage among the different categories. For hourly usage, the results pointed to a peak in app usage in the evening (8:00 pm) for apps targeting mental health problems. The results also showed that mindfulness/meditation apps had two usage peaks: one in the morning (7 am-9 am) and the other in the late evening (10 pm-midnight). In terms of daily usage, the results showed a peak in app usage on Thursday for mindfulness/meditation apps.

Sensitivity Analysis

We conducted a series of Mann-Whitney *U* tests to examine the difference in app open rate, number of sessions, daily minutes of use, and 30-day retention among the top 5 installed apps and the remaining apps per mental health focus and incorporated technique. We found a significant difference in the open rate of mental health apps favoring the top 5 installed apps ($z=1.68$, $P \leq .05$; top 5 installed apps: median 9.0%, IQR 6.9%; remaining apps: $n=54$, median 4.0%, IQR 4.7%). Among these five apps, one incorporated online peer support and three incorporated mood trackers. No other differences were found. A series of Mann-Whitney *U* tests was also conducted to examine whether app usage (app open rates, daily number of sessions, daily minutes of use) in each app category (mental health focus, incorporated technique) differed between apps with or without in-app purchases and no significant differences were found (all $P > .05$).

Figure 4. Hourly usage pattern. Usage is presented by hour out of the total app usage; therefore, the sum of percentages within each category is 100%. Note: a subset of apps for which that data were available is included; “All apps” includes both categories and one app targeting happiness.

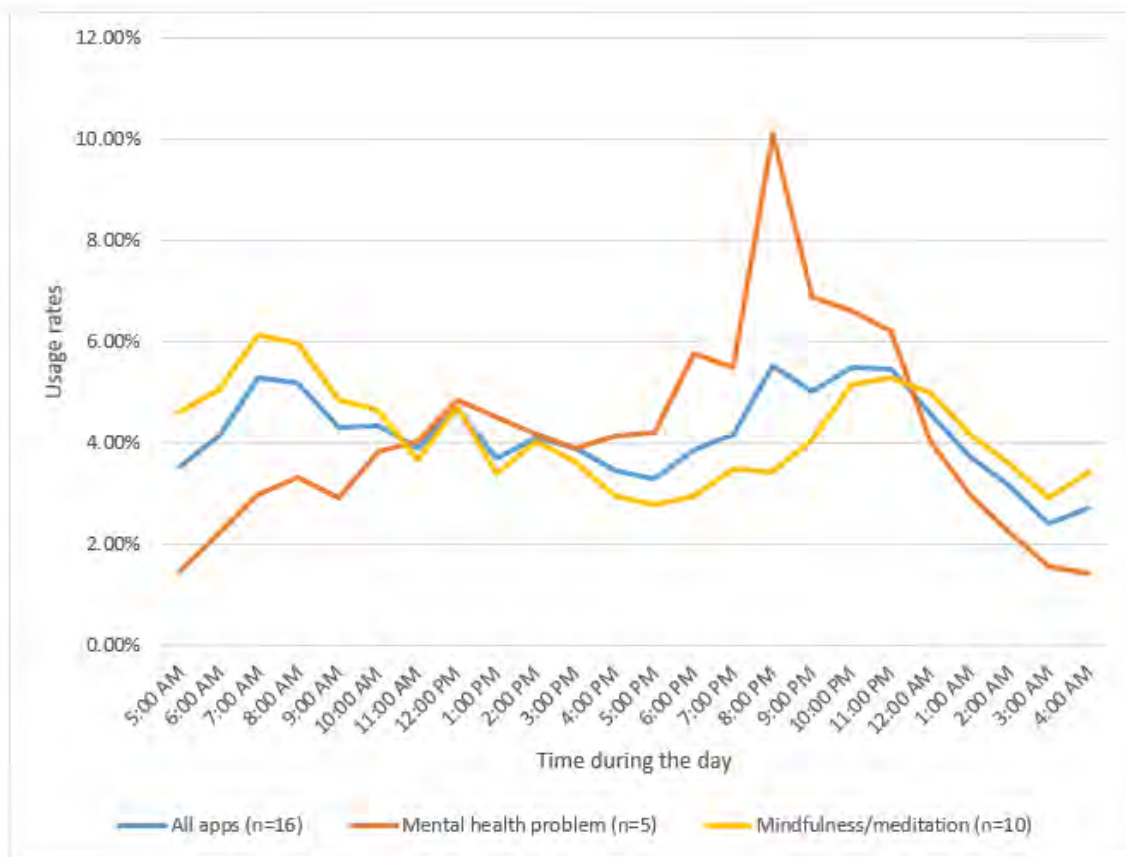


Figure 5. Daily usage pattern. Percentage of app usage is presented by day out of the total app usage; therefore, the sum of percentages within each category is 100%. Note: a subset of apps for which data were available is included; “All apps” includes both categories and one app targeting happiness.



Discussion

Principal Findings

This is the first study to report the usage and retention metrics of a large number of frequently installed, unguided mental health apps as recorded “in the wild” and independent of developer-led data. Based on Google Play Store data (using keyword search terms), there were over 90 million mental health app installs documented by the end of 2018 (ie, reach). Although our findings revealed that daily active users use apps for a significant amount of time during the day (daily usage median of 13.03 minutes), most people with the app installed on their device do not open it in any given day (median open rate of 4.0%). Furthermore, general user retention is poor, with a median 15-day retention of 3.9% and 30-day retention of 3.3%. These findings reflect the lower ranges of real-world retention rates reported in developer-led studies [17-20,22].

Our results also indicate that there are significant differences in app usage and user retention that are associated with the app’s incorporated techniques. Daily minutes of use were significantly higher for mindfulness/meditation (median 21.47) and peer support (median 35.08) apps than for apps incorporating other techniques. Daily open rates were significantly lower for breathing exercise apps (median 1.6%) than for apps incorporating the two techniques with the highest open rates (tracker: median 6.3%; peer support: median 17.0%). User 30-day retention was significantly lower for breathing exercise apps (median 0.0%) than for all other incorporated techniques (mindfulness/meditation: 4.7%; trackers: 6.1%; peer support: 8.9%), except for psychoeducation, which exhibited a pattern similar to the breathing exercise apps at 30-day retention. These patterns could be explained using the notion of *effective engagement* described by Yardley and colleagues [36], wherein there is “sufficient engagement with the intervention to achieve intended outcomes.” From this perspective, it might be that once people acquire the desired skills (breathing exercise) or knowledge (psychoeducation) they no longer use the app, thus affecting the pattern of retention over a longer period. By contrast, mindfulness/meditation apps often include guided meditations designed for repeated use over longer periods of time, while not fostering learning or direct skill acquisition.

Our findings on user retention highlight the low engagement with these apps. Although this warrants a re-evaluation of current engagement and retention strategies, it does not necessarily suggest that these apps are only helpful for a small number of users. First, we do not have data implying that users engage only with one app in the self-management of their states or conditions. However, it is difficult to assume that users are knowledgeable about the different apps available, which apps to use, and when to use them. Although there are some recommender websites [27,29,37] and approaches to help users identify the right apps [38-41], a therapeutic framework that provides guidance to users about how to use the right app at the right time could be useful. For example, in their novel study of IntelliCare—a suite of 13 apps and one Hub app accompanied by 8 weeks of coaching to encourage participants to try the apps recommended to them through the Hub app—Mohr and

colleagues [42] found that 95% of participants eventually downloaded five or more of the IntelliCare apps as part of their therapeutic process. In another study, patients with schizophrenia spectrum disorders received 6 months of treatment that included health technology coaching around the use of three digital tools that were offered to patients based on their needs; 96% of patients rated the program as beneficial [43]. Future studies are needed to examine the feasibility of executing a scalable framework of care in which users receive the right app recommendation at the right time as part of a self-management routine.

Second, user retention patterns might also indicate the low burden associated with app installation (ie, the simplicity of opening the Google Play Store and clicking the app download and installation buttons), which implies that user context, motivation, and ability to engage [44] with these apps were not tested before app installation. The poor active user rates found in our analysis (median open rates of 4%) suggest that the number of app installs available in app stores do not provide a proper estimation of the proportion of users who actually self-manage their state by using the app. These issues further justify a previous call for the development of models to conceptualize the relationships between user state, need, ability, and motivation to engage with early interventions in the digital public space [8]. Although we need to significantly improve our ability to engage users who have made initial attempts at help-seeking, taking a public health engagement approach that is also focused on sustainability represents an important step forward in scaling effective care.

Finally, we identified that the two apps that incorporated peer support as a primary technique had relatively high engagement and retention rates. In our previous work, we defined a program’s *relatability* as “a good representation of a human factor that is easily relatable within the therapeutic context/process” [38]. Relational factors have also been previously acknowledged to nurture a therapeutic alliance with users [45-47], and have demonstrated to be a quality aspect that predicts user engagement with mobile health interventions [28]. Future studies are needed to determine whether technology has a special advantage as an infrastructure that connects between users and results in better engagement rates.

Limitations

This study has several limitations that should be considered. First, because we used an anonymous user panel, we did not have data about how different users use the apps and which parts of the apps were more engaging. The absence of such data means that some apps might have been more engaging due to the characteristics of their users, a phenomenon suggested previously by Ernsting and colleagues [48]. In addition, due to this limitation we were only able to focus on primary incorporated techniques within the apps and not on the way different design features (not deemed to be a primary technique) may have impacted the results. Subsequently, because we were leaning on off-the-shelf programs available to the public, we could not manipulate the programs themselves to account for aspects which lacked variability in our data, such as the impact of theoretical modalities on usage. That is, although our study

advantage is that it enables us to present benchmarks of real-world use independent to trial settings, one advantage of direct experiments is the ability to control participant identity and manipulate intervention modalities and features to identify the group of active components leading to the best outcome (eg, [49]). Such experiments could be also helpful in determining causal relationships between intervention modalities and user behaviors, based on the context of use.

Second, some techniques such as peer support were only incorporated by a small number of highly installed apps (median installation category of 300,000). However, our results did not indicate a significant difference in any incorporated technique in terms of app installs, which suggests that these apps usage patterns go beyond an app's popularity.

Third, because we were led by the available metrics on the platform, we could not examine retention rates after the first 30 days. The retention slope presented a slower decline in app open rates between day 15 and 30 and, based on previous reports, it would be reasonable to assume that there is a continuous usage decline over time (eg, [19,50]), but more studies are needed to determine the magnitude of the decline.

Finally, this study was only based on Android users. Current estimates suggest that the Android market share is approximately

88% of mobile phone users globally [51] and approximately 42.7% of mobile phone users in the United States [52]. Although these data suggest that a sufficient portion of users use the Android operating system, it would be beneficial to validate these results with datasets from the Apple market.

Conclusions

The use of digital platforms that record user traffic “in the wild” enables us to examine patterns of app usage outside of study settings and to assess real-world public engagement. Although we found daily active minutes of use to be relatively high, only a small portion of users actually used popular apps regularly. More studies leveraging different datasets are needed to understand these phenomena. On a broader level, findings point to the importance of the ways we measure, report, and address aspects of user engagement in the real world. It would be helpful to track the context of users who eventually use apps, hopefully through the use of digital footprints, while also tracking the use of multiple apps and websites across times. Obviously, aspects that relate to security and privacy of data have to be addressed. In addition, new studies are needed to better conceptualize our understanding of users' contexts and the ways they search for and engage with beneficial services outside of traditional health care settings.

Acknowledgments

This study was supported by the Donald & Barbara Zucker Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Definition of sham techniques.

[\[PDF File \(Adobe PDF File\)75 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Definition of coded techniques.

[\[PDF File \(Adobe PDF File\)62 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

List of included apps.

[\[PDF File \(Adobe PDF File\)72 KB-Multimedia Appendix 3\]](#)

References

1. Krishna S, Boren S, Balas E. Healthcare via cell phones: a systematic review. *Telemed J E Health* 2009 Apr;15(3):231-240. [doi: [10.1089/tmj.2008.0099](#)] [Medline: [19382860](#)]
2. Norman GJ, Zabinski MF, Adams MA, Rosenberg DE, Yaroch AL, Atienza AA. A review of eHealth interventions for physical activity and dietary behavior change. *Am J Prev Med* 2007 Oct;33(4):336-345 [FREE Full text] [doi: [10.1016/j.amepre.2007.05.007](#)] [Medline: [17888860](#)]
3. Naslund J, Marsch L, McHugo G, Bartels S. Emerging mHealth and eHealth interventions for serious mental illness: a review of the literature. *J Ment Health* 2015;24(5):321-332 [FREE Full text] [doi: [10.3109/09638237.2015.1019054](#)] [Medline: [26017625](#)]
4. Pew Research Center. 2019 Jun 12. Mobile fact sheet URL:<http://www.pewinternet.org/fact-sheet/mobile/> [accessed 2019-01-07] [WebCite Cache ID 75F2i2NKz]

5. WalkerSands Communications. 2019. How technology is expanding the scope of online commerce beyond retail URL:<https://www.walkersands.com/resources/the-future-of-retail-2018/> [accessed 2019-01-07] [WebCite Cache ID 75F2nwClg]
6. Anthes E. Mental health: there's an app for that. *Nature* 2016 Apr 7;532(7597):20-23. [doi: [10.1038/532020a](https://doi.org/10.1038/532020a)] [Medline: [27078548](https://pubmed.ncbi.nlm.nih.gov/27078548/)]
7. Kazdin A. Addressing the treatment gap: a key challenge for extending evidence-based psychosocial interventions. *Behav Res Ther* 2017 Jan;88:7-18. [doi: [10.1016/j.brat.2016.06.004](https://doi.org/10.1016/j.brat.2016.06.004)] [Medline: [28110678](https://pubmed.ncbi.nlm.nih.gov/28110678/)]
8. Baumel A, Baker J, Birnbaum ML, Christensen H, De Choudhury M, Mohr DC, et al. Summary of key issues raised in the Technology for Early Awareness of Addiction and Mental Illness (TEAAM-I) meeting. *Psychiatr Serv* 2018 May 01;69(5):590-592. [doi: [10.1176/appi.ps.201700270](https://doi.org/10.1176/appi.ps.201700270)] [Medline: [29334875](https://pubmed.ncbi.nlm.nih.gov/29334875/)]
9. Eysenbach G. The law of attrition. *J Med Internet Res* 2005;7(1):e11 [FREE Full text] [doi: [10.2196/jmir.7.1.e11](https://doi.org/10.2196/jmir.7.1.e11)] [Medline: [15829473](https://pubmed.ncbi.nlm.nih.gov/15829473/)]
10. Christensen H, Mackinnon A. The law of attrition revisited. *J Med Internet Res* 2006;8(3):e20. [Medline: [17032636](https://pubmed.ncbi.nlm.nih.gov/17032636/)]
11. Baumeister R, Vohs K. Self-regulation, ego depletion, and motivation. *Social Pers Psych Compass* 2007 Nov;1(1):115-128. [doi: [10.1111/j.1751-9004.2007.00001.x](https://doi.org/10.1111/j.1751-9004.2007.00001.x)]
12. Kohl LF, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. *J Med Internet Res* 2013;15(7):e146 [FREE Full text] [doi: [10.2196/jmir.2665](https://doi.org/10.2196/jmir.2665)] [Medline: [23859884](https://pubmed.ncbi.nlm.nih.gov/23859884/)]
13. Mohr DC, Weingardt KR, Reddy M, Schueller SM. Three problems with current digital mental health research...and three things we can do about them. *Psychiatr Serv* 2017 May 01;68(5):427-429. [doi: [10.1176/appi.ps.201600541](https://doi.org/10.1176/appi.ps.201600541)] [Medline: [28412890](https://pubmed.ncbi.nlm.nih.gov/28412890/)]
14. Day J, Sanders M. Do prents benefit from help when completing a self-guided parenting program online? A randomized controlled trial comparing Triple P Online with and without telephone support. *Behav Ther* 2018;49(6):1020-1038.
15. Glasgow RE. eHealth evaluation and dissemination research. *Am J Prev Med* 2007 May;32(5 Suppl):S119-S126. [doi: [10.1016/j.amepre.2007.01.023](https://doi.org/10.1016/j.amepre.2007.01.023)] [Medline: [17466816](https://pubmed.ncbi.nlm.nih.gov/17466816/)]
16. Farvolden P, Denisoff E, Selby P, Bagby RM, Rudy L. Usage and longitudinal effectiveness of a Web-based self-help cognitive behavioral therapy program for panic disorder. *J Med Internet Res* 2005;7(1):e7 [FREE Full text] [doi: [10.2196/jmir.7.1.e7](https://doi.org/10.2196/jmir.7.1.e7)] [Medline: [15829479](https://pubmed.ncbi.nlm.nih.gov/15829479/)]
17. Christensen H, Griffiths KM, Korten AE, Brittcliffe K, Groves C. A comparison of changes in anxiety and depression symptoms of spontaneous users and trial participants of a cognitive behavior therapy website. *J Med Internet Res* 2004 Dec 22;6(4):e46 [FREE Full text] [doi: [10.2196/jmir.6.4.e46](https://doi.org/10.2196/jmir.6.4.e46)] [Medline: [15631970](https://pubmed.ncbi.nlm.nih.gov/15631970/)]
18. Fleming T, Bavin L, Lucassen M, Stasiak K, Hopkins S, Merry S. Beyond the trial: systematic review of real-world uptake and engagement with digital self-help interventions for depression, low mood, or anxiety. *J Med Internet Res* 2018 Jun 06;20(6):e199 [FREE Full text] [doi: [10.2196/jmir.9275](https://doi.org/10.2196/jmir.9275)] [Medline: [29875089](https://pubmed.ncbi.nlm.nih.gov/29875089/)]
19. Owen JE, Jaworski BK, Kuhn E, Makin-Byrd KN, Ramsey KM, Hoffman JE. mHealth in the wild: using novel data to examine the reach, use, and impact of PTSD Coach. *JMIR Ment Health* 2015;2(1):e7 [FREE Full text] [doi: [10.2196/mental.3935](https://doi.org/10.2196/mental.3935)] [Medline: [26543913](https://pubmed.ncbi.nlm.nih.gov/26543913/)]
20. Carpenter J, Crutchley P, Zilca RD, Schwartz HA, Smith LK, Cobb AM, et al. Seeing the "big" picture: big data methods for exploring relationships between usage, language, and outcome in internet intervention data. *J Med Internet Res* 2016 Aug 31;18(8):e241 [FREE Full text] [doi: [10.2196/jmir.5725](https://doi.org/10.2196/jmir.5725)] [Medline: [27580524](https://pubmed.ncbi.nlm.nih.gov/27580524/)]
21. Ebert D, Baumeister H. Internet-based self-help interventions for depression in routine care. *JAMA Psychiatry* 2017 Aug 01;74(8):852-853. [doi: [10.1001/jamapsychiatry.2017.1394](https://doi.org/10.1001/jamapsychiatry.2017.1394)] [Medline: [28636717](https://pubmed.ncbi.nlm.nih.gov/28636717/)]
22. Christensen H, Griffiths K, Farrer L. Adherence in internet interventions for anxiety and depression. *J Med Internet Res* 2009 Apr 24;11(2):e13. [doi: [10.2196/jmir.1194](https://doi.org/10.2196/jmir.1194)] [Medline: [19403466](https://pubmed.ncbi.nlm.nih.gov/19403466/)]
23. Moller A, Merchant G, Conroy D, West R, Hekler E, Kugler KC, et al. Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. *J Behav Med* 2017 Feb;40(1):85-98 [FREE Full text] [doi: [10.1007/s10865-016-9818-7](https://doi.org/10.1007/s10865-016-9818-7)] [Medline: [28058516](https://pubmed.ncbi.nlm.nih.gov/28058516/)]
24. Baumel A, Yom-Tov E. Predicting user adherence to behavioral eHealth interventions in the real world: examining which aspects of intervention design matter most. *Transl Behav Med* 2018 Sep 08;8(5):793-798. [doi: [10.1093/tbm/ibx037](https://doi.org/10.1093/tbm/ibx037)] [Medline: [29471424](https://pubmed.ncbi.nlm.nih.gov/29471424/)]
25. Demyttenaere K, Bruffaerts R, Posada-Villa J, Gasquet I, Kovess V, Lepine JP, WHO World Mental Health Survey Consortium. Prevalence, severity, and unmet need for treatment of mental disorders in the World Health Organization World Mental Health Surveys. *JAMA* 2004 Jun 02;291(21):2581-2590. [doi: [10.1001/jama.291.21.2581](https://doi.org/10.1001/jama.291.21.2581)] [Medline: [15173149](https://pubmed.ncbi.nlm.nih.gov/15173149/)]
26. World Health Organization. The World Health Report 2001. Mental Health: New Understanding, New Hope. Geneva: World Health Organization; 2001.
27. MindTools.io. URL:<https://mindtools.io/> [accessed 2018-06-24] [WebCite Cache ID 70Pg0KTXJ]
28. Baumel A, Kane J. Examining predictors of real-world user engagement with self-guided eHealth interventions: analysis of mobile apps and websites using a novel dataset. *J Med Internet Res* 2018 Dec 14;20(12):e11491 [FREE Full text] [doi: [10.2196/11491](https://doi.org/10.2196/11491)] [Medline: [30552077](https://pubmed.ncbi.nlm.nih.gov/30552077/)]
29. PsyberGuide. URL:<http://psyberguide.org/> [accessed 2015-04-10] [WebCite Cache ID 6Xh99eQmZ]

30. Baumel A, Birnbaum ML, Sucala M. A systematic review and taxonomy of published quality criteria related to the evaluation of user-facing eHealth programs. *J Med Syst* 2017;41(8):128. [doi: [10.1007/s10916-017-0776-6](https://doi.org/10.1007/s10916-017-0776-6)] [Medline: [28735372](#)]
31. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006 Jan;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
32. SimilarWeb. 2018. URL:<https://www.similarweb.com/> [accessed 2018-06-18] [WebCite Cache ID 70GZ6eAt5]
33. European Commission. Data protection URL:https://ec.europa.eu/info/law/law-topic/data-protection_en [accessed 2018-06-18] [WebCite Cache ID 70GY1yLYp]
34. Oath. URL:<https://www.oath.com/> [accessed 2018-10-14] [WebCite Cache ID 73A1knb5w]
35. Carlson KD, Herdman AO. Understanding the impact of convergent validity on research results. *Organ Res Methods* 2010 Dec 30;15(1):17-32. [doi: [10.1177/1094428110392383](https://doi.org/10.1177/1094428110392383)]
36. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. *Am J Prev Med* 2016 Nov;51(5):833-842. [doi: [10.1016/j.amepre.2016.06.015](https://doi.org/10.1016/j.amepre.2016.06.015)] [Medline: [27745683](#)]
37. ADAA Mental Health Apps rates. URL:<https://www.adaa.org/finding-help/mobile-apps> [accessed 2016-11-01] [WebCite Cache ID 6lhtKNvkc]
38. Baumel A, Faber K, Mathur N, Kane JM, Muench F. Enlight: a comprehensive quality and therapeutic potential evaluation tool for mobile and web-based eHealth interventions. *J Med Internet Res* 2017 Mar 21;19(3):e82 [FREE Full text] [doi: [10.2196/jmir.7270](https://doi.org/10.2196/jmir.7270)] [Medline: [28325712](#)]
39. Baumel A, Muench F. Heuristic evaluation of eHealth interventions: establishing standards that relate to the therapeutic process perspective. *JMIR Ment Health* 2016 Jan 13;3(1):e5 [FREE Full text] [doi: [10.2196/mental.4563](https://doi.org/10.2196/mental.4563)] [Medline: [26764209](#)]
40. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR Mhealth Uhealth* 2015;3(1):e27 [FREE Full text] [doi: [10.2196/mhealth.3422](https://doi.org/10.2196/mhealth.3422)] [Medline: [25760773](#)]
41. Torous JB, Chan SR, Gipson SY, Kim JW, Nguyen T, Luo J, et al. A hierarchical framework for evaluation and informed decision making regarding smartphone apps for clinical care. *Psychiatr Serv* 2018 May 01;69(5):498-500. [doi: [10.1176/appi.ps.201700423](https://doi.org/10.1176/appi.ps.201700423)] [Medline: [29446337](#)]
42. Mohr DC, Tomasino KN, Lattie EG, Palac HL, Kwasny MJ, Weingardt K, et al. IntelliCare:an eclectic, skills-based app suite for the treatment of depression and anxiety. *J Med Internet Res* 2017 Jan 05;19(1):e10 [FREE Full text] [doi: [10.2196/jmir.6645](https://doi.org/10.2196/jmir.6645)] [Medline: [28057609](#)]
43. Baumel A, Correll CU, Hauser M, Brunette M, Rotondi A, Ben-Zeev D, et al. Health technology intervention after hospitalization for schizophrenia: service utilization and user satisfaction. *Psychiatr Serv* 2016 Jun 1;67(9):1035-1038. [doi: [10.1176/appi.ps.201500317](https://doi.org/10.1176/appi.ps.201500317)] [Medline: [27247171](#)]
44. Fogg B. A behavior model for persuasive design. 2009 Presented at: Persuasive '09 4th international Conference on Persuasive Technology; April 26-29, 2009; Claremont, CA. [doi: [10.1145/1541948.1541999](https://doi.org/10.1145/1541948.1541999)]
45. Cavanagh K, Millings A. (Inter)personal computing: the role of the therapeutic relationship in e-mental health. *J Contemp Psychother* 2013 Jul 17;43(4):197-206. [doi: [10.1007/s10879-013-9242-z](https://doi.org/10.1007/s10879-013-9242-z)]
46. Holter MT, Johansen A, Brendryen H. How a fully automated eHealth program simulates three therapeutic processes: a case study. *J Med Internet Res* 2016 Jun 28;18(6):e176 [FREE Full text] [doi: [10.2196/jmir.5415](https://doi.org/10.2196/jmir.5415)] [Medline: [27354373](#)]
47. Barazzone N, Cavanagh K, Richards DA. Computerized cognitive behavioural therapy and the therapeutic alliance: a qualitative enquiry. *Br J Clin Psychol* 2012 Nov;51(4):396-417. [doi: [10.1111/j.2044-8260.2012.02035.x](https://doi.org/10.1111/j.2044-8260.2012.02035.x)] [Medline: [23078210](#)]
48. Ernsting C, Dombrowski S, Oedekoven M, O Sullivan JL, Kanzler M, Kuhlmeier A, et al. Using smartphones and health apps to change and manage health behaviors: a population-based survey. *J Med Internet Res* 2017 Apr 05;19(4):e101 [FREE Full text] [doi: [10.2196/jmir.6838](https://doi.org/10.2196/jmir.6838)] [Medline: [28381394](#)]
49. Collins L, Murphy S, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. *Am J Prev Med* 2007 May;32(5 Suppl):S112-S118 [FREE Full text] [doi: [10.1016/j.amepre.2007.01.022](https://doi.org/10.1016/j.amepre.2007.01.022)] [Medline: [17466815](#)]
50. Cheung K, Ling W, Karr C, Weingardt K, Schueller S, Mohr D. Evaluation of a recommender app for apps for the treatment of depression and anxiety: an analysis of longitudinal user engagement. *J Am Med Inform Assoc* 2018 Aug 01;25(8):955-962 [FREE Full text] [doi: [10.1093/jamia/ocy023](https://doi.org/10.1093/jamia/ocy023)] [Medline: [29659857](#)]
51. Statista. Global mobile OS market share in sales to end users from 1st quarter 2009 to 2nd quarter 2018 URL:<https://www.statista.com/statistics/266136/global-market-share-held-by-smartphone-operating-systems/> [accessed 2019-01-06] [WebCite Cache ID 75DpPpzz6]
52. Statcounter GlobalStats. Mobile operating system market share United States of America URL:<http://gs.statcounter.com/os-market-share/mobile/united-states-of-america> [accessed 2019-01-06] [WebCite Cache ID 75Dr29YOS]

Abbreviations

ANOVA: analysis of variance

IQR: interquartile range

Edited by G Eysenbach; submitted 02.05.19; peer-reviewed by A Rotondi, B Chaudhry, L Vizer, MS Aslam, L Bonet; comments to author 31.05.19; revised version received 01.07.19; accepted 19.07.19; published 15.09.19

Please cite as:

Baumel A, Muench F, Edan S, Kane JM

Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis

J Med Internet Res 2019;21(9):e14567

URL: <http://www.jmir.org/2019/9/e14567/>

doi: [10.2196/14567](https://doi.org/10.2196/14567)

PMID:

©Amit Baumel, Frederick Muench, Stav Edan, John M Kane. Originally published in the Journal of Medical Internet Research (<http://www.jmir.org>), 15.09.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on <http://www.jmir.org/>, as well as this copyright and license information must be included.

EXHIBIT 15

INFORMATION ANALYSIS

Analysis of the Webometric Indicators of the Main Websites that Aggregate Multithematic Scientific Information

Yu. M. Brumshteyn* and E. Yu. Vas'kovskii**

Astrakhan State University, Astrakhan, Russia

**e-mail: brum2003@mail.ru*

***e-mail: vaskovskiy_evgeniy@mail.ru*

Received August 2, 2017

Abstract—A rationale is provided for the importance of assessing the availability, volumes, and relevance of scientific and technical information (STI) on content-aggregator websites. It is suggested that STI be classified into two groups: scientific information and engineering/technical information. Based on certain categorization principles described in this article, content-aggregator websites are divided into 33 categories. A discussion is given of the possibilities and drawbacks of the webometrics methodology that is used in Russia to evaluate websites with scientific content. A new methodology is proposed for evaluating the volumes and relevance of information on aggregator websites of scientific content. The proposed methodology has been tested on the websites of several leading foreign and Russian organizations aggregating multithematic scientific information. A discussion is given about the place of these websites in the global and Russian scientific infospace, the composition and sources of the scientific information aggregated on the websites, the terms of access to this information, and whether and how these websites use systems for calculating scientometric indicators. The webometric indicators are presented and discussed for all the websites included in the study.

Keywords: scientific information, content-aggregator websites, classification of aggregators, information and communication technologies, webometric indicators, web analytics, information volumes, web-traffic indicators, relevance of information, user behavior

DOI: 10.3103/S0005105517060048

INTRODUCTION

Increasing the efficiency of the production and use of scientific and technical information (STI), including through its aggregation on websites with easy access to it, and enhancing Russia's integration into the international scientific infospace are necessary conditions for Russia's successful social and economic development. Meeting the latter condition has, in turn, two main implications: (1) increasing the popularity, availability, and relevance of Russian STI in the international infospace and (2) improving the accessibility and efficiency of the use of foreign STI by Russian researchers, including through awareness of the nomenclature and capabilities of foreign databases (DBs) and other STI repositories.

The improved production and use of STI in Russia is extremely important for addressing the following practical tasks: facilitation of the intellectual development of the country as a whole, as well as individual regions, organizations, and researchers; support of the effective functioning/development of socio-economic and socio-technical systems at different hierarchical levels; improvement of the information infrastructure of society; improvement of the information

support for scientific and engineering activities through the support of access to operational and retrospective information; support of scientific communications, including international ones; and development of scientific schools and new areas of research and development [1–5]. However, given the rapid growth in the amounts of information posted on the Internet (including in electronic scientific libraries), a real issue occurs regarding the reliability and objectivity of this information [6].

The availability of STI to users has several aspects: the principal possibility of access to the necessary scientific information from a given country or organization; free access for users to information units in stored in repositories; the presence of access barriers due to high payments for access to scientific information relative to the personal incomes of researchers or financial opportunities of their organizations (this also applies to academic users); the usability of the websites, first of all, their data-retrieval systems (DRSs); the “open hours” of the websites; and the duration for processing requests in the DRSs of the websites [7–13].

In the Soviet Union, the following organizations played a key role in the accumulation and, partly, dis-

semination of scientific information: libraries of various types [1, 14], including scientific, engineering, departmental, educational, etc., ones; territorial and departmental STI centers (STICs); STI depository organizations including departmental ones; and organizations that aggregate data on STI of Soviet and foreign origins, primarily the All-Russian (formerly, All-Soviet) Institute for Scientific and Technical Information (VINITI) [15].

Various technologies have been used to provide access to information: (1) temporary lending of materials (books, journals, copies of descriptions of inventions and author's certificates, information sheets, etc.) from libraries and STICs to members for use in reading rooms and, in some cases, for taking home; (2) making copies of available materials (entirely or fragmentally) to paper, film, or microfiches, both in the presence of users and by mailing copies ordered remotely; in some cases, these copies were used to form library funds; (3) obtaining special publications (or copies thereof) requested by researchers through interlibrary loans, for temporary use; (4) prompt distribution to users of copies of tables of contents for new issues of periodicals, including inaccessible Soviet and foreign scientific journals; (5) centralized planning/management of the printing of scientific literature by publishers; (6) maintenance of various types of bibliographic indexes in libraries and STICs on the materials available in their funds; indexes (keys) to catalogs; chronicles of journal articles; lists of new releases of books, journals, etc.; making it convenient for those interested to find the necessary materials in the vast funds of these organizations; (7) periodic exhibitions of new literature received by the funds of individual libraries; all-Union exhibitions of new literature released by publishers; thematic exhibitions during scientific congresses, etc.; (8) the VINITI abstract journal (AJ), which consisted of various thematic series [15], played a key role in the aggregation and promotion of information materials (especially foreign ones). At least a part of the series was available in major libraries and in other organizations. Based on the brief abstracts published in the AJ (sometimes author's abstracts or simply titles of papers), researchers could order copies of the necessary information, for which they either paid themselves or requested support from their host institutions; and (9) some scientific journals (primarily, academic ones) mailed the authors copies of their articles published in these journals.

The development of information and telecommunication technologies (ITCT) has fundamentally changed the situation. (A) It is now critical to perform on-the-fly monitoring and analysis of the global STI flow [16], as soon as it appears. (B) The functions of the modern STI system in Russia [17] have expanded, including the support of innovation processes, training of specialists, etc. (C) Websites, as well as their

DRSs and Internet search engines, now play the key role in the accumulation and accessibility of STI for potential users in Russia and abroad [2, 8, 18, 19]. (D) The availability and relevance of scientific information has increased due to the widespread use of mobile devices, including ipads, smartphones, laptops, etc. (E) Broad development has occurred in the automated tools of scientometric analysis of scientific information, including indexes of individual researchers as well as journals, organizations, countries, etc. [20]. (F) A sharp increase has been observed in the attention to the integration of Russia into the international infospace. This is especially true of publications by Russian scientists in those scientific journals, etc. whose materials are indexed in international citation systems and scientometric indicators recognized by Russia's Higher Attestation Commission (HAC) [20]. The term *recognized* hereafter means that publications in these journals, etc. are noted separately by the HAC during the defenses of dissertations, etc. At the same time, based on the HAC's resolution no. 1-pl/1 of June 15, 2017, the status of these publications is higher than that of works published in Russian journals included in the lists of journals recommended by the HAC for the publication of the results of candidate and doctoral dissertations.

In this regard, we should note the following: (1) a substantial portion of STI, especially that related to promotion and advertising, is hosted on websites without any quality control and reliability checks. Sometimes this information is heavily biased, being aimed only at promoting specific inventions, goods, or information and other services; (2) when users make queries on Internet search engines the engines often give too many links, often with unsatisfactory sorting by relevance. Reformulation of search queries to improve the quality of the search results does not always give satisfactory outcomes. The possibilities for improving the quality of processing of search queries are, in practice, rather limited, even if we use advanced search tools, intellectualization tools, etc; and (3) the information hosted on STI aggregator websites [1, 17, 19] is monitored and indexed. This allows users to select materials by using the internal DRSs of these websites, which usually provide both simple and advanced search capabilities. However, even these search engines often give many irrelevant links for adequate search queries. Nevertheless, researchers tend to look for scientific information mainly on aggregator websites. Therefore, it is currently important to evaluate the usability of these websites [13] and their relevance for users, based on the various web analytics tools (including web traffic counters) and altmetrics methods. Webometric indicators (WMIs) of websites, including their dynamics, are fundamentally important for the analysis and management of STI flows, for determining the efficiency of scientific institutions

and universities, for evaluating and managing the usability of their sites, for building (calibrating) mathematical models of STI flows, etc. [21–29]. Publications exist that are specifically devoted to studying WMIs of the websites of Russian scientific institutions [28, 30].

THE GENERAL ISSUES OF THE ACCESSIBILITY AND QUALITY OF SCIENTIFIC INFORMATION

Most of the existing scientific publications and normative documents are focused on STI. However, for analysis of the webometric indicators of content-aggregator websites, we suggest dividing STI into scientific and engineering/technical (including industry-related) information. There are often no clear boundaries between these two groups; thus, they can be posted on both same and different Internet resources. We only consider aggregator websites of scientific information, although some of these sites also post engineering and technical information.

Scientific information (SI) includes various types (categories) of documents, which are aggregated on websites in varying degrees and from different sources. We emphasize that SI is not confined to published scientific articles and monographs; it also includes other types of documents, such as information on upcoming scientific events, personal information of scientists and their achievements, information about scientific schools [4], etc. An important role in the availability of SI is played by Internet service tools. They can be integrated from DBs of aggregator websites or be hosted on separate websites (i.e., use external DBs).

The sources of aggregated SI are, first of all, publications in various scientific periodicals, conference proceedings, collective papers, monographs (including collective ones), preprints, etc. It is assumed that these materials are subject to some kind of editorial control over the quality and reliability of the information. The information space is protected from poor-quality (low-level) SI through targeted measures, such as the rigorous selection that journals undergo to be accepted in the leading international citation systems; assessment of scientometric indicators; and the limited number of sources from which one selects publications in *Google Scholar* (GS) and other specialized SI DRSs; certain requirements for publications included in Russia's HAC lists. We also note the decision of the Russian Scientific Citation Index (RSCI) of April 29, 2017 to exclude 344 editions from among those indexed by this resource.

In the typical cases, aggregator organizations receive scientific publications in the electronic format, with the necessary hypertext markup, which allows these organizations to minimize the labor costs of adding the new materials to the DBs.

Internet sites can post both copies of the same materials and links to the web repositories containing the originals.

The first option leads to information redundancy. However, it also increases the reliability of storage of materials, thus increasing the likelihood that they will be found by Internet search engines and the internal DRS of the aggregator websites. Specifically, materials published in scientific journals of many Russian universities are posted both on the websites of these institutions and on *eLIBRARY.ru*, *CyberLeninka*, etc.

The second option for posting SI excludes duplication of materials and distortion of bibliographic information when copying. However, this method is more vulnerable to information security threats, especially in the absence of mirrors for the servers that contain repositories of the original materials.

One way to improve the potential availability of information is, in the general case, to translate SI into other languages and perform targeted popularization (information promotion) of research results [31–33]. In the Soviet Union, foreign documents were made more accessible by translating the most important ones into Russian; information about these translations was disseminated using subject bibliographic indexes, which were sent to libraries and STICs.

Today, professional Russian translations are posted on Internet sites together with the original documents. However, Internet sites, as rule, have no special indexes of translated materials. The capabilities of machine translation are limited in terms of quality due to the ambiguity of scientific terminology and complexity of grammar structures.

Why do Internet sites publish relatively few Russian translations of foreign scientific documents? We can give the following reasons: the low demand for such translations, especially in narrow subject fields; the relatively high level of English proficiency among scientists, who often have a good command of terminology in their fields; the availability of translation software tools (including web-based translators) and electronic dictionaries with word combinations that are typically used in various subject areas and sources; and partial compensation of deficiencies in the stylistics of electronic translations through the professional knowledge of scientists who specialize in the relevant subject areas.

We will emphasize some of the factors that are important for the subsequent presentation.

The existing approaches for evaluating the relevance of electronic SI are focused on posting this information on content-aggregator websites.

The actual relevance of Russian-language scientific materials is higher than the number of references in scientific publications that are taken into account by the RSCI and other citation and scientometric systems.

Extended relevance assessments for individual Russian-language scientific documents can potentially be made using the RSCI altmetrics (such as the number of views, downloads, and inclusions in collections). However, the RSCI does not use these parameters in calculating their indicators for the following reasons: (1) comprehensive techniques for calculating SI relevance are not available yet and (2) altmetrics gives more opportunities for falsification than counting the number of references in scientific publications.

The altmetrics for websites differs from that for articles. The use of these indicators has been vigorously discussed in scientific literature [24–26], including such indicators as *website traffic*, which is treated as a characteristic of Internet relevance [23]. However, the traffic on a particular website (including its dynamics over time) can only be determined by system administrators by installing special software on servers [22] or by connecting external webometrics tools [22, 28, 30]. One would usually find no traffic data for individual websites on the Internet. Therefore, the information on changes in website traffic (daily, weekly, or monthly) can be used in decision making only by managers of those sites.

Most scientific materials worldwide are now published in English. The number of English-language materials published by Russian authors in foreign journals is growing rapidly. Many such articles also appear in Russian journals. Many scientific materials are now published in other languages, including in Chinese, Japanese, Korean, French, and Spanish. They are made accessible to Russian-speaking users largely through the publication of their English-language abstracts and sometimes through the use of translation software tools and electronic dictionaries. Russian users can read some scientific publications in the Belarusian and (with some reservations) Ukrainian languages, without translation.

Editors of Russian scientific journals usually welcome articles written in English because they increase the likelihood that the journals will be included in international citation and scientometric systems recognized by Russia's HAC. Most Russian authors find it considerably more difficult to prepare an article in English than in Russian. Russian-speaking users have less access to these English-language articles than to Russian-language ones, leading to a lower number of references to such publications in Russian literature.

Russian-language publications are made available to English-speaking scientists by posting English-language abstracts or versions of articles on websites and through the efforts of MAIK/Nauka Interperiodika to publish English translation of some Russian scientific journals. The last option (Russian publication + English publication) is most appealing for authors: the articles are available for both Russian- and foreign-speaking scientists. This potentially leads to an increase in the number of references to these articles

[34], improves their altmetrics, and expands the personal contacts of researchers.

THE STATUS OF RUSSIAN SCIENTIFIC JOURNALS AND CITATION–SCIENTOMETRIC SYSTEMS FROM THE PERSPECTIVE OF THE HAC WITH THE RUSSIAN MINISTRY OF EDUCATION AND SCIENCE

The HAC uses two categories of peer-reviewed scientific journals recommended for publishing the results of candidate and doctoral dissertations: journals included in the international citation and scientometric systems recognized by the HAC (Russian journals are given in **C1** list) and peer-reviewed Russian scientific journals in which researchers should publish the main scientific results of their candidate and doctoral dissertations (**C2** list).

As of July 1, 2017, the HAC recognized the following international citation and scientometric systems: *Web of Science*, *Scopus*, *Springer*, *ZbMath*, *MathSciNet*, *GeoRef*, *PubMed*, *Agris*, and *Astrophysics Data System (ADS)*.¹ The first three resources are universal (multithematic) and the rest of them should be considered specialized. To be included in these systems, publishers of scientific journals must submit applications, undergo rigorous selection procedures to ensure compliance with the norms of publishing ethics, the quality and relevance of published materials, and the consistency of the content of the published articles with their titles and abstracts and with the declared thematic areas of the publisher. As of the end of June 2017, the **C1** and **C2** lists included 1003 and 2143 Russian journals, respectively.

Categorization of SI aggregator websites can be based on different approaches [16]. To divide the aggregator websites into categories by their characteristics, we use the following terminology:

- T1, citation and scientometric systems;
- T2, bibliographic DBs (contain bibliographic data only);
- T3, research abstract DBs, which contain abstracts of materials from scientific journals and other editions;
- T4, full-text DBs of scientific materials;
- T5, search systems for scientific articles and other materials;
- T6, search, information, and information–analytical systems for scientific journals and websites;
- T7, service tools for working with personal data of scientists. Some of the Internet resources considered below have several functionalities out of those listed above.

¹ According to the resolution of June 15, 2017 no.1-pl/1, the HAC of Russia will not recognize the *Agris* and *ADS* systems as of January 1, 2018.

SI is published, to some extent or another, on various aggregator websites. However, there is a considerable difference between the popularity of these sites and their relevance. In this respect, the competition between websites is also important, especially if users can access the same materials on different resources. Specifically, the websites of many Russian universities and research institutes post not only the texts of scientific articles published by these institutions but also materials related to dissertation defenses. However, in most cases, users can find the information they need (especially on dissertations) on these websites only if they have a very clear understanding of what to look for and where.

We identify 33 categories of SI aggregator websites, based on the following criteria: the purpose and functionality of the websites; the nature of the information published on the websites; the necessity and sufficiency of all the categories; and the convenience of analysis (comparison) between foreign and Russian resources.

K1. The websites of international multithematic (universal) citation and scientometric systems recognized by the HAC of Russia.

K2. The websites of international multithematic (universal) SI repositories that are not recognized by the HAC of Russia, including resources containing assessments of scientometric indicators.

K3. Major Russian Internet resources of universal SI.

K4. International thematically specialized citation and scientometric systems recognized by the HAC of Russia.

K5. International thematically specialized SI repositories that are not recognized by the HAC of Russia.

K6. Foreign websites specifically designed for information support of national publications.

K7. Internet resources with catalogs of dissertations, including their extended abstracts and full texts.

K8. Information and information-analytical systems on scientific journals.

K9. Foreign websites of individual journals or groups of journals with a common publisher.

K10. Russian websites of individual journals or groups of journals with a common publisher.

K11. The websites for assigning DOI to articles and ORCID to authors.

K12. Thematically specialized Russian SI resources (repositories) that are not classified in other categories.

K13. The websites of international and foreign scientific and scientific/technical organizations, their associations, scientific societies.

K14. The websites of Russian scientific and scientific/technical organizations, their associations, scientific societies.

K15. The websites of foreign comprehensive/multidiscipline educational institutions.

K16. The websites of Russian comprehensive/multidiscipline educational institutions.

K17. The websites of foreign specialized educational institutions.

K18. The websites of Russian specialized educational institutions.

K19. Universal and specialized SI search systems.

(a) Only structures collections of hyperlinks to Internet resources, sometimes in combination with information on the content of these resources.

(b) Systems with their own DBs (index catalogs), including the addresses of DBs with scientific materials on external resources and the volumes and update schedules of these DBs.

(c) Systems that automatically collect and aggregate, upon user requests, information from all the available Internet sources.

K20. Annotated catalogs of information resources, including on service tools.

K21. Systems for detecting replicated texts (instances of plagiarism or incorrect citation) in scientific articles and other materials posted on websites in electronic form.

K22. The websites of traditional libraries in foreign countries and in Russia.

K23. The websites of publishing houses that specialize in scientific literature.

K24. Electronic library systems of publishing houses.

K25. The websites of electronic libraries that are not classified in the above categories.

K26. Specialized sites with summary information on upcoming or past scientific conferences, workshops, and other scientific events.

K27. Internet resources with information about scientific organizations and their ratings.

K28. Resources with personal information about individual scientists and teams of scientists.

K29. The websites of professional social networks and other resources designed to facilitate scientific communications.

K30. Websites with normative documents on scientific activities and scientific publishing ethics.

K31. The websites of Russian and foreign foundations that support scientific research, organization of scientific events, etc.

K32. Websites associated with research personnel training, employment opportunities for scientists, internship programs, etc.

K33. Websites containing summary statistics, analytical overviews, and other information on accomplished scientific results and expected outcomes of scientific activities.

Since the above classification applies to SI only, it does not reflect the groups of aggregator websites with patent information, with information about software, DBs, etc. This kind of information on Russian inventions and developments is available, e.g., on the website of the Federal Institute of Industrial Property (FIIP) of Russia.

The issues relating to the support for information connections between the listed categories of Internet resources (including through the use of hyperlinks) are highly relevant. However, this range of issues requires separate consideration.

It is highly important that STI aggregator organizations should carefully choose sources of aggregated information, taking its reliability, relevance, quality, and coverage into account; should base their work on the principles and practical technologies of aggregating information from various sources [9, 33] avoiding unjustified duplication of materials on the same resources; use thematic classifications and assign adequate descriptors to the stored materials; and ensure information security for individual information materials and aggregator websites in general.

Due to size limitations, in this article we confine our analysis only to the webometric indicators of the websites in categories K1–K3.

RESEARCH METHODOLOGY

We selected content-aggregator websites from three categories under consideration (K1–K3) based on their popularity in the international and Russian information space. We searched for website addresses using search engines (machines) on the websites: <http://www.google.ru> and <http://www.yandex.ru> and for information on aggregator organizations using <https://ru.wikipedia.org>.

We obtained information about the activities and sizes of the resources both from their own websites and from external sources. Note that the size data for the same aggregator websites may differ considerable in different sources, for reasons other than different dates of retrieval.

The need for an objective assessment of website webometrics leads to the development of web analytics tools [21], including the development and use of webometric indexes. The latter are, in fact, integral indicators that allow one to summarize information from a set of specific WMIs, which, in turn, can be assessed in the following ways:

(v1) From the data communicated by system administrators or authorized representatives of organizations to an organization that concentrates this information. These data are collected by software located on the websites (including traffic counters [22]) and/or by web traffic analysis tools such as Google Analytics or Yandex.Metrics. In these cases, it is possible to estimate/compare the traffic on different

groups of webpages or even on individual pages. We emphasize that this approach works only in a very limited number of cases because most organizations are not interested in communicating their WMI data, comparing them with other organizations, and publishing this information in public domain.

In rare instances, the homepages of aggregator websites may indicate the current number of visitors or the accumulated number of views over a certain period of time. It would make sense to install separate counts for visitors from the following categories: from home organizations within the local computer network (system administrators can easily estimate this number); users outside the home organizations but from the home locality (home city, etc.); from other localities in Russia; foreign users (preferably with a breakdown by country).

(v2) Information published on special webometric sites [30], which cover a range of specifically monitored institutional websites.

(v3) Information about websites from the main Internet search engines. This approach is applied, e.g., in [28, 30] in a comparative analysis of the websites of academic institutions in Russia. The results of monitoring these resources are posted on the website <http://webometrix.ru>. However, the authors of [28] noted that the data from different Internet search engines differ substantially. Note that the processes of improving these systems may also lead to changes in the WMIs of aggregator websites. As a consequence, data for different periods may not be fully comparable.

To derive website rankings, the authors of [28] calculated an index by the algorithm from <http://w.ict.nsc.ru/ranking/>. The algorithm uses specific WMIs such as (a) the arithmetic average of the number of external links to the website in three systems (*Yandex*, *Google*, and *Bing*), not taking the relative popularities of these search engines into account; (b) the arithmetic average of the number of pages on the website (in the same three systems); (c) the arithmetic average of the number of PDF, DOC, and PPT files (in total), averaged only for *Google* and *Yandex*; (d) citation indexes from the *Yandex* and *Google Scholar* citation index systems. We note that in the literature there is an ongoing debate about the use of *Google Scholar* for such purposes [35]. All the four parameters (a)–(d) were then used to rank the organizations. For each of the organizations, the ranks in all the four rankings were summed. By summing the ranks (and not the primary data), one can work with heterogeneous indicators. However, attaching, by default, the same importance to each of the four parameters is, in our view, logically vulnerable.

Therefore, in [28] five indicators were taken into account in deriving the website rankings.

(1) The size of the website, as estimated by the number of pages.

(2) The visibility, as estimated by the number of incoming links to the website, as detected by Internet search engines.

(3) The altmetrics, as understood in [28] as visibility in the networks for professional communication and cooperation of scientists.

(4) The number of full-text files, as detected by Internet search engines (machines).

(5) Scientificity, as estimated in [28] as the number of links to the website and materials on it, as detected by *Google Scholar*.

If we talk about the quality of scientific materials, it seems that we should also consider the number of citations for each object in *Google Scholar*. To calculate the webometric index of a particular site, the authors of [28] assigned weight coefficients to the five listed indicators: 0.1, 0.25, 0.25, 0.1, and 0.30, respectively. This weighting system was apparently taken from www.webometrics.info, consistent with the methodology of the Cybernetic Laboratory at the Center for Scientific Information and Documentation of the Spanish National Research Council. The technique in [28] has, in our view, the following disadvantages: for 1 and 2 ([28, p. 4]), the authors overlooked the *Yandex* and *Rambler* systems despite their relative popularity in Russia; for 3 [28], they overlooked the social networks *Vkontakte* (and, with some reservations, *Odnoklassniki*), which are popular in Russia and are also used by Russian scientists.

For Russia, it makes sense to develop national webometric indexes [36]. In practice, the first such indexes were devised primarily for the websites of Russian scientific institutions by different centers independently [28, 30, 37].

(v4) Information about website indicators obtained from *Google Scholar* queries [28, 30, 37]. It is believed that these data show the quality of materials posted in the public domain because *Google Scholar* only indexes individual objects. Most of these materials are archives of scientific journals, proceedings of scientific conferences, and scientific papers. However, *Google Scholar* may assign no indexes at all to the contents of some websites, especially those that require a login and a password to access the resources.

(v5) Data collected by special software (including public and paid tools), which enable remote analysis of website indicators.

(v6) Data on website hyperlinks that are collected by specially designed crawler robots [37], which may specialize in certain thematic areas [38].

In this article, we applied a WMI assessment methodology using approaches based on v5 and v4. It is largely similar to the one we described previously in [39]. Below, we reproduce this methodology in part to ensure that readers understand the information in the tables (which use generally accepted English terms and

their translations). We assessed the WMIs of content-aggregator websites in five areas.

(1) Website availability (*Ping*) was estimated by the average time (AT) for opening the start page in seconds. To this end, we used a service that checks the load speed for webpages from different locations around the world (<http://sitespeed.me>).

(2) Website content parameters (*Content*): (a) The size of the website content: this parameter was determined using the website's address system, based on the number of addresses of static or dynamically generated files (*Count*, *URLs*). We used the following parameters in this subgroup: the number of website pages (*text/html*); the number of image files (*image*) loaded into the website's file system and used on the site; and the number of files posted on the site (*application*). These designations are consistent with the Internet Media Types specification, i.e., the standard file-type identifiers (*Content-Type*).

The values of these parameters were obtained from standard reports generated by Xenu's Link Sleuth software (<http://xenu-link-sleuth.en.softonic.com>). It checks websites for hyperlinks by URLs located in the following objects: tags <a>, image tags, plug-ins, local image maps, style sheets, scripts, and Java applets. Xenu's Link Sleuth browses the links to webpages and also checks all the links on these pages. This algorithm potentially allows one to check the entire website, but it may be too time consuming on large websites.

(b) Website size in megabytes (*Size*, *Mb*) was also determined using Xenu's Link Sleuth.

(3) Website visibility (*Visibility*) was assessed according to the number of links to it on other websites, i.e., by the number of incoming links. To determine visibility, we used the service hosted at <https://majestic.com>.

(4) The convenience of working with information posted on a website was assessed using the following parameters. (a) The number of internal links (NIL) on the website (φ) was estimated by the formula

$$\varphi = \sum_{j=1}^m LIn_j, \quad (1)$$

where m is the number of internal URLs on the website and LIn_j is the number of links to the j th internal URL on the website. In so doing, we considered only the URLs of files of the text/html type, i.e., with an indication of webpages.

(b) The number of outgoing links (NOL) from the website was taken into account using two indicators: (b1) the number of unique links to webpages (the number of outgoing URLs on the website), denoted as Ω ; (b2) the total number of outgoing links (links to

outgoing URLs on the website), denoted as Ψ . The NOL was estimated by the formula

$$\Psi = \sum_{i=1}^n LIn_i, \quad (2)$$

where n is the number of outgoing URLs on the website and LIn_i is the number of links on the website to the i th outgoing URL.

The NIL and NOL were estimated on the basis of standard reports generated by Xenu's Link Sleuth.

In addition to the above absolute indicators for links, we also calculated relative ones, normalized by the number of webpages. This allowed us to compare the analyzed characteristics for websites of different sizes.

(5) Website relevance was assessed according to three indicators: (a) the number of unique visitors per month; (b) the number of visits per month; and (c) average visit duration (AVD), i.e., the average duration of the visitor's work with the website during one access session. For indicators a and b, the month was a period of 30 days preceding the date of assessment.

Indicators a and b (in area 5) were estimated using a service hosted on <http://pr-cy.ru/>. This service automatically checks open statistics counters and derives the basic traffic parameters. First, it searches for an open Yandex.Metrika counter. If it contains no data about the site, the service checks the LiveInternet resource. If it finds nothing in these sources, then the system <http://pr-cy.ru> makes a site traffic forecast on the basis of a rating system based on the ratio between the traffic on specific sites (the total number and frequency of visits) and other websites. These estimates use the Alexa Rank resource at <http://www.alexa.com>. The lower the Alexa Rank is, the more popular the Internet resource is. The service hosted on <http://pr-cy.ru> implements an algorithm allowing one to convert the forecasted Alexa Rank indicator into approximate forecast traffic data for a given website.

The AVD indicator (c) was estimated using the service hosted on <https://www.similarweb.com>.

The total duration of the works of users with a content-aggregator website can be estimated as a product of the AVD multiplied by the total number of visits. However, this indicator should be treated with caution because users can keep the page of an aggregator website open in a minimized window and use it only from time to time. A deeper analysis (available to system administrators only) may involve the determination of average indicators of user activity per site visit: the time interval between performing any actions, including scrolling the pages with the results of data-retrieval systems and the texts of open files, adding materials to personal collections, etc.; the number of completed requests; the number of opened and/or downloaded files.

We also estimated the number of website objects indexed in *Google Scholar*.

Since the content-aggregator websites discussed in this article differ greatly among themselves, we do not propose any universal (integral) assessment criteria.

Another factor that complicates data analysis for content-aggregator websites is that some resources have no websites of their own but are hosted on the websites of other organizations or their departments. In this case, commonly used programs do not allow external users to obtain metrics for the page groups corresponding to the aggregator websites. However, the indicators of these groups (and even individual pages) can be obtained by system administrators of the aggregator websites by using counters for individual pages [22] and by connecting to the websites of external web analytics tools [21, 22].

All the information in Tables 1–3 relates to the period from July 1, 2017 to July 25, 2017 (we could not obtain data for a shorter period due to the high labor intensity). Since the data changes over time, the analysis of WMI dynamics can present a separate area of research. Table 1 compares WMIs obtained by different approaches.

We can draw the following conclusions from Table 1: (1) the results obtained using different approaches differ significantly; (2) different Internet search engines (machines) generate significantly different results for the same sites; and (3) the use of *Google Scholar* to estimate the number of links on some websites is unproductive.

Moreover, in comparison with the existing approaches that are described at the beginning of this section our methodology is more informative, especially in website relevance assessments.

FOREIGN MULTITHEMATIC CITATION AND SCIENTOMETRIC SYSTEMS RECOGNIZED BY THE HAC OF RUSSIA

In this and subsequent sections, we first describe the resources in the corresponding categories and then present tables with WMIs for these resources.

Scopus (www.scopus.com), a DB of research abstracts and scientometric data. This resource indexes over 18 500 titles of scientific/technical and medical journals published by approximately 5000 organizations.

To access most of the SI posted on the resource, interested organizations must sign agreements and make payments to the resource owner. Therefore, in Russia, the information on *Scopus* is available in full only for major universities, research institutes, and major libraries. The indicators below relate to the resource on www.scopus.com in general.

Web of Science (WoS), a resource hosted on www.webofknowledge.com: the Science Citation Index Expanded (SCI-EXPANDED), information on natural and selected applied sciences; the Social

Table 1. Comparison of webometric indicators obtained by different approaches

| Resource | Authors' methodology | | Approaches applied in [28] | | | |
|--------------------------------|----------------------|------------|----------------------------|---------------|----------|-----------------------|
| | | | <i>Yandex</i> | <i>Google</i> | | <i>Google Scholar</i> |
| | NIL* | WPN* | WPN | NIL | WPN | WPN |
| <i>www.scopus.com</i> | 7433278 | 1 | 1 | 1050000 | 1 | 0 |
| <i>www.webofknowledge.com</i> | 1632266 | 1 | 1 | 2740000 | 2610 | 0 |
| <i>www.springer.com</i> | 16152325 | >1000000** | 209000 | 29800000 | 1100000 | 89 |
| <i>www.springeropen.com</i> | 281455 | 56945 | 1000 | 309000 | 20000 | 18 |
| <i>doaj.org</i> | 3581433 | 59 | 23000 | 189000 | 514000 | 0 |
| <i>indexcopernicus.com</i> | 986019 | 138 | 0 | 81200 | 4 | 22 |
| <i>www.proquest.com</i> | 3542085 | 5962 | 2000 | 3720000 | 5260 | 0 |
| <i>www.researchbib.com</i> | 104156 | 15 | 19 | 86700 | 187 | 0 |
| <i>www.science.gov</i> | 171925 | 176 | 246000 | 8670000 | 695000 | 1 |
| <i>www.sciencedirect.com</i> | 25891818 | 79 | 384000 | 4740000 | 14700000 | 9550000 |
| <i>www.hub.sciverse.com</i> | 7336 | 2 | 1 | 49600 | 1 | 0 |
| <i>www.scienceresearch.com</i> | 8324 | 34 | 0 | 3430000 | 26 | 0 |
| <i>cogprints.org</i> | 69007 | 198 | 8000 | 119000 | 109000 | 1750 |
| <i>sci-hub.io</i> | 96094 | 3 | 2 | 1320000 | 2 | 0 |
| <i>libgen.io</i> | 323244 | 189 | 2 | 3190000 | 3670 | 6 |
| <i>elibrary.ru</i> | 191615 | 4 | 3000000 | 767000 | 4440000 | 1670000 |
| <i>cyberleninka.ru</i> | 1188351 | 1673138 | 2000000 | 175000 | 1300000 | 1270000 |
| <i>www2.viniti.ru</i> | 55214 | 294 | 383 | 23100 | 322 | 1 |
| <i>sci-lib.com</i> | 751643 | 348907 | 11000 | 2220000 | 585000 | 0 |
| <i>gbu.bookchamber.ru</i> | 1082 | 29 | 37 | 113 | 46 | 0 |
| <i>sciencepublic.ru</i> | 479 | 60 | 5 | 76900 | 746 | 50 |
| <i>arbicon.ru</i> | 210371 | 607 | 3000 | 6050 | 18300 | 156 |
| <i>archive.neicon.ru</i> | 218698 | 62779 | 0 | 100 | 64100 | 110 |

* NIL is the number of incoming links to the website; WPN is the number of webpages on the website (in the authors' methodology, the WPN is the sum of the indicators *Text/Html* and *Application* in the table with even numbers).

** Due to the large amounts of data that are processed, we could not obtain a more accurate value.

Sciences Citation Index (SSCI); and the Arts&Humanities Citation Index (A&HCI). The information in these indexes relates to the period since 1991. These resources do not contain full-text files but include bibliographic references to publications.

The website of the HAC presents abbreviated designations for the different *WoS* indexes in the section entitled Information on List C1 Journals. Here we decipher them, based on the data from different sources: (a) SCI–Expanded (Science Citation Index Expanded); (b) ESCI–Emerging Science Citation Index, a part of the Web of Science Core Collection; according to the information on <http://www.open-science.in.ua/ru/esci.html>, ESCI accepts regional-level journals covering local issues; (c) BIOSIS stands for BIOSIS Citation Index; and (d) ZR stands for Zoological Record. The above-mentioned abbreviations SSCI and A&HCI are not included in the HAC's List C1 at all.

List C1 does not explicitly mention the abbreviation RSCI (Russian Science Citation Index), which reflects core RSCI scientific journals. However, some information about these journals is given at https://ru.wikipedia.org/wiki/Web_of_Science. Note that apart from RSCI, *WoS* contains information of the same status about other groups of journals, e.g., KCI, the Korean Journal Database.

Here we should also mention the *WoS* web analytics website *Clarivate Analytics*: <https://clarivate.com/products/web-of-science/>, which also has a Russian version (<http://wokinfo.com/russian/>).

Springer (www.springer.com), a multithematic DB of one of the world's largest media holdings. The website claims to keep record of articles from 2900 journals and of approximately 280000 books on sale. The website has its own data-retrieval system, which does not require user registration. The resource uses a convenient system of thematic rubrics.

Table 2. Indicators for K1 aggregator websites (part 1)

| N | Resource | AT, s | Count, URLs | | | Size, MB | GS |
|---|--|-------|-------------|-------|-------------|----------|----|
| | | | text/html | image | application | | |
| 1 | <i>Scopus</i> | 1.17 | 1 | 0 | 0 | 0 | 0 |
| 2 | <i>Web of Science</i> | 0.83 | 1 | 0 | 0 | 0 | 0 |
| 3 | <i>Clarivate Analytics</i> https://clarivate.com/products/web-of-science/ | 5.83 | 8 | 96 | 8 | 1.03 | 0 |
| | <i>Clarivate Analytics</i> http://wokinfo.com/russian/ | 0.37 | 660 | 187 | 135 | 250 | 0 |
| 4 | <i>Springer</i> | 0.33 | >1000000* | ** | ** | ** | 89 |
| 5 | <i>SpringerOpen</i> | 0.44 | 7983 | 310 | 48962 | 70766 | 18 |

SG is the number of objects (materials) indexed by *Google Scholar*.

* Due to the large amounts of data that are processed, we could not obtain a more accurate value.

** Due to the large amounts of data that are processed, we could not calculate this value.

Table 3. Indicators for K1 aggregator websites (part 2)

| Resource | | Absolute number of links | | | | Unique visitors per month | Visits per month | AVD, hh:mm:ss |
|----------|--|--------------------------|--------|----------|--------|---------------------------|------------------|---------------|
| | | in | ext | out | | | | |
| | | | | Ω | Ψ | | | |
| 1 | <i>Scopus</i> | 7433278 | 0 | 0 | 0 | 346906 | 1387620 | 00:05:23 |
| 2 | <i>Web of Science</i> | 1632266 | 0 | 0 | 0 | 497556 | 1990230 | 00:07:25 |
| 3 | <i>Clarivate Analytics</i> https://clarivate.com/products/web-of-science/ | 65590 | 40 | 254 | 1702 | *** | *** | *** |
| | <i>Clarivate Analytics</i> http://wokinfo.com/russian/ | 180098 | 6680 | 614 | 6165 | *** | *** | *** |
| 4 | <i>Springer</i> | 16152325 | ** | ** | ** | 1918445 | 7673790 | 00:02:37 |
| 5 | <i>SpringerOpen</i> | 281455 | 180455 | 49665 | 196623 | 126233 | 504930 | 00:01:40 |

** Due to the large amounts of data that are processed, we could not calculate this value.

*** The values can be measured only for the website as a whole.

We separately note the website *SpringerOpen* (www.springeropen.com), which concentrates articles from a limited number of open-access journals. This website also has a convenient system of rubrics.

Tables 2 and 3 show the WMIs for content-aggregator websites in category K1.

Based on Table 2, we can conclude that the described tools do not allow one to estimate the sizes of most of the aggregator websites. The reasons are as follows: the websites require a login and a password; the software tools operate for too long on large websites. Moreover, we should keep in mind that SI DBs used by these Internet resources and by those considered below can be located outside the aggregator websites.

We can draw the following conclusions from Table 3:

(1) Many of the listed sites have very high WMIs because they are global in nature, serving English-speaking users.

(2) Our methodology cannot be used to estimate some of the WMIs of these websites due to the high time costs of measuring the WMIs.

(3) The AVD for *Scopus* and *Web of Science* are much higher than those for *Springer* and *SpringerOpen*. However, the number of unique visitors for *Springer* is much larger than for *Scopus* and *Web of Science* combined. Such a situation may be due not only to the differences between the amounts of information in paid and free access on the websites being compared but also to the different popularity of the websites among users.

Table 4. Indicators for K2 aggregator websites (part 1)

| N | Resource | AT, s | Count, URLs | | | Size, Mb | SG |
|----|--|-------|-------------|-------|-------------|----------|---------|
| | | | text/html | image | application | | |
| 1 | <i>Directory of Open Access Journals</i> | 0.93 | 47 | 55 | 12 | 5.5 | 0 |
| 2 | <i>Wilson Company</i> | 5.6 | 2 | 8 | 0 | 0.039 | 0 |
| 3 | <i>INSPEC</i> | 0.4 | 3 | 12 | 0 | 0.189 | 0 |
| 4 | <i>Index Copernicus</i> | 0.88 | 136 | 72 | 2 | 21.1 | 22 |
| 5 | <i>ProQuest</i> | 0.7 | 5962 | 2 | 0 | 226 | 0 |
| 6 | <i>ResearchBib</i> | 2.4 | 15 | 6 | 0 | 0.2 | 0 |
| 7 | <i>Science.gov</i> | 1.11 | 118 | 210 | 58 | 69 | 1 |
| 8 | <i>www.sciencedirect.com</i> | 0.7 | 79 | 0 | 0 | 2 | 9550000 |
| 9 | <i>www.hub.sciverse.com</i> | 0.54 | 2 | 1 | 0 | 0.005 | 0 |
| 10 | <i>www.scienceresearch.com</i> | 0.33 | 33 | 89 | 1 | 5.2 | 0 |
| 11 | <i>LAR</i> | 3.78 | 116 | 29 | 0 | 4.3 | 0 |
| 12 | <i>Cogprints</i> | 0.38 | 195 | 21 | 3 | 12.5 | 1750 |
| 13 | <i>SCI-HUB</i> | 0.19 | 2 | 17 | 1 | 0.55 | 0 |
| 14 | <i>Library Genesis</i> | 0.67 | 160 | 9 | 29 | 195 | 6 |

(4) Unfortunately, open-access software does not allow one to obtain statistics for these resources by users from Russia only.

FOREIGN MULTITHEMATIC SI REPOSITORIES NOT RECOGNIZED BY THE HAC OF RUSSIA

Here we present only those resources that can be regarded as major ones; this selection is, to some extent, subjective.

DOAJ (*Directory of Open Access Journals*, <https://doaj.org/>). To ensure that materials published in journals (including Russian ones) are accepted for coverage by this resource, the publishers should take additional action and make payments. Therefore, some Russian journals, including those listed by the HAC, do not seek inclusion in this resource.

Recently, DOI registration agencies began to take money for assigning DOI names to articles in some journals. However, it is a common practice abroad to publish scientific articles in *DOAJ* (and these articles receive a DOI). It is believed that potentially, *DOAJ* inclusion may considerably increase the number of links to the articles and, hence, improve scientometric indicators for journals and authors. Here, we emphasize that the inclusion of published works in *DOAJ* is only possible for open-access journals.

Wilson Company (<http://www.hwwilson.com/>). Based on the data at http://www.nlr.ru/res/inv/ic_www/cat_show.php?p=2&rid=77, *Wilson Company* offers 39 full-text and 25 bibliographic DBs (including 14 research abstract DBs). Therefore, we consider this resource within category K2. However, an attempt to open the website <http://www.hwwilson.com/> automat-

ically redirects the user to <https://www.ebsco.com/products/research-databases/h-w-wilson-databases>.

INSPEC (<http://www.theiet.org/resources/inspec>). This resource contains abstracts from over 3500 journals; it includes over 8 million entries and covers 2000 conference proceedings and a large number of books, dissertations, and reports.

Index Copernicus (<http://indexcopernicus.com>). This Polish scientometric DB is considered as an international platform for cooperation between scientists. This resource can post scientific materials by initiative of individual authors and journals.

ProQuest (<http://www.proquest.com/>). This resource is intended to support libraries, scientific publishers, and distributors of scientific literature. The DB of this resource contains scientific articles, ebooks, and over 3 million dissertations. The resource has an interface in several languages but does not have a Russian version.

ResearchBib (<https://www.researchbib.com/>). This resource has the subtitle *Academic Resource Index*. It contains archives of articles in several languages and indexes scientific journals, conference proceedings, employment opportunities for scientists, and internship programs.

Science.gov (<https://www.science.gov/>). An American website that publishes governmental SI and results of researches supported from the federal budget. This resource works with 60 DBs and 2200 websites. It provides access to over 200 million pages. This website can be considered as an umbrella resource with respect to the ones displayed on it.

Elsevier (<http://www.sciencedirect.com/>). The multithematic resources of *Science Direct*.

Table 5. Indicators for K2 aggregator websites (part 2)

| N | Resource | Absolute number of links | | | | Unique visitors per month | Visits per month | AVD, hh:mm:ss |
|----|-----------------------------------|--------------------------|--------|----------|--------|---------------------------|------------------|---------------|
| | | in | ext | out | | | | |
| | | | | Ω | Ψ | | | |
| 1 | Directory of Open Access Journals | 3581433 | 554 | 471 | 763 | 79272 | 317100 | 00:04:06 |
| 2 | Wilson Company | 8474 | 1 | 111 | 111 | **** | **** | **** |
| 3 | INSPEC | 21753 | 2 | 9 | 9 | **** | **** | **** |
| 4 | Index Copernicus | 986019 | 2575 | 22 | 350 | 40499 | 162000 | 00:05:12 |
| 5 | ProQuest | 3542085 | 129795 | 115329 | 80253 | 543553 | 2174220 | 00:04:01 |
| 6 | ResearchBib | 104156 | 175 | 65 | 205 | 26750 | 107010 | 00:02:53 |
| 7 | Science.gov | 171925 | 571 | 1237 | 1483 | 63343 | 253380 | 00:00:48 |
| 8 | www.sciencedirect.com | 25891818 | 130 | 203 | 776 | 2347970 | 9391890 | 00:03:45 |
| 9 | www.hub.sciverse.com | 7336 | 1 | 4 | 4 | 8574 | 34290 | 00:00:10 |
| 10 | www.scienceresearch.com | 8324 | 288 | 132 | 195 | 9763 | 39060 | 00:05:49 |
| 11 | LAR | 303 | 3923 | 42 | 315 | **** | **** | **** |
| 12 | Cogprints | 69007 | 1785 | 39 | 544 | 20388 | 81540 | 00:00:54 |
| 12 | SCI-HUB | 96094 | 2 | 6 | 6 | 307902 | 1231620 | 00:05:50 |
| 13 | Library Genesis | 323244 | 1818 | 1768 | 4452 | 721243 | 2974800 | 00:07:37 |

**** The WMI software tools we used have no access to the statistics for this website.

Elsevier (<http://www.hub.sciverse.com/>). The multithematic resources of *Science Hub Diverse*.

The resource <http://www.scienceresearch.com>. This resource can also be considered as a search system on publications in *Science Research*.

LAR (<http://www.thelar.org/2015-01-30-20-08-08/detailed-information-about-lar-journals>). A multithematic DB on scientific journals and articles therein, including open-access ones. Table 4 presents data on this resource as a whole, i.e., for <http://www.thelar.org>.

Cogprints (<http://cogprints.org>). A resource with open-access materials.

SCI-HUB (<https://sci-hub.io>). Based on the information posted on this resource, it provides access to 62 million scientific articles. The creator of this resource fundamentally rejects intellectual property rights and copyrights for scientific and educational information because paid-access article are unaffordable, at least for students. A specific feature of its internal data-retrieval system is the possibility to type in a URL, PMI/DOI, or search string. The website also offers a special extension for *Google Chrome*, which can be downloaded and installed from the site. The resource works with the pirate library *Library Genesis*.

Library Genesis (<http://libgen.io>). This website contains a search system and an online repository that provide free access to various materials, not necessarily scientific ones.

We do not include *Google Scholar* in category K2 because we treat it as a specialized search system for scientific materials that are available on the Internet.

All the indicators for category K2 websites are grouped in Tables 4 and 5.

The modest estimates for the sizes of most of the resources included in Table 4 suggest that they work with DBs located on external sources. We should point out that SG detects a large number of objects in only one case (www.sciencedirect.com).

We can draw the following conclusions from the results presented in Table 5:

(1) The AVD values differ considerably among the resources.

(2) The AVD for open-access resources is smaller than the corresponding indicator for *Springer* and is on par with *Scopus* and *WoS*.

(3) The greatest AVD values are observed for pirate resources.

(4) The higher number of visits per month among the category K2 websites is recorded for www.sciencedirect.com (its indicator is even higher than for *Springer*).

(5) However, the visits-per-month indicators for the pirate resources are also very high, largely exceeding, e.g., those for *DOAJ*.

MAJOR RUSSIAN MULTITHEMATIC SI RESOURCES

We emphasize that many of the category K3 resources in this section can be classified in other categories. However, due to their special role or func-

tions, we chose to consider these resources in this section.

1. The host website of the Russian resources *National Electronic Library* (NEL) [40] and the RSCI is *eLIBRARY.ru* (elibrary.ru). This is a multifunctional resource that includes: DBs of scientific materials; an internal data-retrieval system; personal data on researchers; and calculated results for scientometric indicators. For a substantial part of the publications in the NEL, this resource provides access to detailed information on the published materials: it either displays the information on the website or redirects the user via hyperlinks to external Internet resources. If a user wants to gain access to the full texts of published articles or address some other issues, he/she should register in the system. The advantage of this website is the detailed scientometric indicators calculated by the RSCI for individual journals, for groups of journals issued by certain institutions, and for individual authors. The website also provides altmetrics for individual publications but not for journals. The website's start page displays the number of users currently working at the website (usually from a few thousand to the first several tens of thousands of people). Recently, the RSCI has been actively engaged in improving the quality of the content [41] by rejecting journals that publish materials without peer review and low-quality papers. For most Russian-speaking researchers, this website now plays a crucial role, especially in solving the problems of applied scientometrics [42].

2. The scientific electronic library *CyberLeninka* (www.cyberleninka.ru). This resource provides free access to scientific articles published in Russia and CIS countries, including in Russian journals on the HAC's lists C1 and C2. The website has a well-developed (multifunctional) data-retrieval system. A substantial number of materials on this resource duplicates those posted on elibrary.ru.

Note that Internet search engines (including *Google*) usually find materials posted on elibrary.ru in *CyberLeninka* as well.

3. The VINITI website with STI catalogs (www2.viniti.ru). This resource provides searches for various types of information. Its DBs contain mainly research abstracts. A part of the materials can only be accessed via a login/password, which are assigned in the VINITI reading rooms or under contracts for paid remote access. This is one of the few websites that display the processing times of search queries.

4. The website of the Russian National Public Library for Science and Technology hosts several catalogs, including an electronic catalog at <http://library2.gpntb.ru/>.

5. The website of the Large Scientific Library (LSL) <http://sci-lib.com/>. This resource contains mainly information on natural sciences and technology. This resource has a web forum to discuss scientific problems.

6. The open-access resource *Electronic Chronicles of the Russian Book Chamber* (gbu.bookchamber.ru). It displays not only books but also journal/newspaper articles, extended abstracts of dissertations, reviews, printed music, and cartographic publications.

7. The website <http://sciencepublic.ru/>, designed by the International Research Federation SciencePublic, contains two main resources: (A) *Science Public Index* (<http://direct.sciencepublic.ru/>), which is positioned as a scientometric DB and (B) *Science Public Direct* (<http://direct.sciencepublic.ru/>), positioned as an open-access DB of scientific materials. Some of the Russian paid editions claim that they not only post materials to *Science Public Direct* but also assign IDSP numbers to all materials posted to *SP Library*, which is described at id.sciencepublic.ru as an international scientific electronic library. In this respect, we should note that *SP Library* is also identified by Internet search engines as the website of the Singapore Polytechnic Library (<https://library.sp.edu.sg/>).

8. The interregional analytical list of articles *MARS* (<http://arbicon.ru>), a project of the Association of Regional Library Consortia (ARBICON), an interdepartmental interregional library network that unites 220 libraries of various systems and departments. The DB includes 2.1 million articles from more than 2000 Russian journals, beginning from 2001.

9. The scientific journal archive *NEICON* (<http://archive.neicon.ru>) provides access to archives of scientific journals from western (foreign) publishers, for a fairly long time period. However, it is difficult to estimate the extent of duplication of materials relative to *eLIBRARY.ru*.

The WMIs for the category K3 websites are summarized in Tables 6 and 7.

The results summarized in Table 6 show that (1) the amounts of information on the first two websites are rather large and comparable to the resources considered in the previous sections (categories K1 and K2); (2) our methods did not allow us to estimate the proportion of Russian/English-language materials on these websites (the English materials may, however, represent not only articles by foreign authors but also those written by Russian researchers and translations of Russian-language articles); (3) the software used in this study did not allow us to analyze the information volumes for the VINITI website. The high AT indicator (opening time of the start page) for this website is also noteworthy. This time may be considerably shorter for users from Russia only.

We can draw the following conclusions from Table 7:

(1) The most relevant resources for Russian-speaking users are the first two websites in Tables 6 and 7.

(2) In terms of the number of unique visitors per month, *CyberLeninka* is far ahead of the NEL/RSCI website (elibrary.ru).

Table 6. Indicators for K3 aggregator websites (part 1)

| N | Resource | AT, s | Count, URLs | | | Size, MB | SG |
|---|----------------------------------|-------|-------------|-------|-------------|----------|---------|
| | | | text/html | image | application | | |
| 1 | <i>eLIBRARY.ru</i> | 1.28 | 4 | 9 | 0 | 0.106 | 1670000 |
| 2 | <i>CyberLeninka</i> | 0.72 | 947921 | 2091 | 725217 | 7260891 | 1270000 |
| 3 | <i>VINITI</i> | 14.94 | 107 | 107 | 187 | 252 | 1 |
| 4 | <i>http://library2.gpntb.ru</i> | 0.66 | 2156 | 10 | 0 | 73 | 0 |
| 5 | <i>Large Scientific Library</i> | 0.37 | 348907 | 71 | 0 | 12492 | 0 |
| 6 | <i>RBC Electronic Chronicles</i> | 0.51 | 27 | 156 | 2 | 3.7 | 0 |
| 7 | <i>sciencepublic.ru</i> | 2.43 | 57 | 21 | 3 | 7.6 | 50 |
| 8 | <i>MARS</i> | 2.54 | 548 | 439 | 59 | 237 | 156 |
| 9 | <i>NEICON Archive</i> | 0.86 | 62779 | 16 | 0 | 2811 | 110 |

Table 7. Indicators for K3 aggregator websites (part 2)

| N | Resource | Absolute number of links | | | | Unique visitors per month | Visits per month | AVD, hh:mm:ss |
|---|----------------------------------|--------------------------|----------|----------|---------|---------------------------|------------------|---------------|
| | | In | Ext | Out | | | | |
| | | | | Ω | Ψ | | | |
| 1 | <i>eLIBRARY.ru</i> | 191615 | 4 | 0 | 0 | 266829 | 1067310 | 00:06:23 |
| 2 | <i>CyberLeninka</i> | 1188351 | 3073859 | 4646 | 4742015 | 366829 | 1467330 | 00:02:31 |
| 3 | <i>VINITI</i> | 55214 | 3941 | 203 | 684 | 8736 | 34950 | 00:2:46 |
| 4 | <i>http://library2.gpntb.ru</i> | 16 | 9580 | 135 | 1571 | 21691 | 86760 | 00:01:56 |
| 5 | <i>Large Scientific Library</i> | 751643 | 71509137 | 73916 | 6442776 | 4906 | 13048 | 00:00:42 |
| 6 | <i>RBC Electronic Chronicles</i> | 1082 | 43 | 4 | 8 | 13111 | 52440 | 00:02:23 |
| 7 | <i>sciencepublic.ru</i> | 479 | 666 | 37 | 292 | 3708 | 14820 | 00:02:54 |
| 8 | <i>MARS</i> | 210371 | 4493 | 851 | 2607 | 9051 | 36210 | 00:03:05 |
| 9 | <i>NEICON Archive</i> | 218698 | 1404408 | 14 | 391933 | 11315 | 21605 | 00:00:58 |

(3) However, these figures for both sites are thousands of visitors per month, rather than millions, compared with some of the foreign sites considered above. This situation reflects the much smaller number of Russian-speaking researchers worldwide in comparison with English-speaking ones. However, if we normalize the number of site visitors by estimates for the numbers of English- and Russian-speaking researchers, then at least the first two Russian K3 websites (from Tables 6 and 7) even have some advantage over the functionally similar foreign aggregator websites.

(4) For the websites of NEL/RSCI and *CyberLeninka*, the number of visits is much larger than that of unique visitors. Hence (also taking into account the AVD), we can conclude that many researchers use these resources on a regular basis;

(5) For the NEL/RSCI website (*elibrary.ru*), the number of unique visitors per month (266 829) was slightly less than a half of the number of authors of publications indexed by the RSCI (578 495). What is the share of the authors among unique site visitors? It is impossible to determine this figure using the soft-

ware that is available to external users of this website. However, considering that the RSCI includes in its number of authors those authors who are deceased, we can conclude that this website is rather popular among living authors.

(6) The question of whether seasonal factors (holidays/vacations of university teachers and researchers in Russia) influences the traffic on *eLIBRARY.ru*, *CyberLeninka*, etc. requires additional study.

(7) The AVD indicator of *eLIBRARY.ru* is more than twice as high as those of *CyberLeninka* and the VINITI website (the AVDs for the latter two websites are approximately the same). This may indicate a greater amount of work performed by users on *eLIBRARY.ru* during one visit to the website than on *CyberLeninka* and the VINITI website.

(8) The AVD indicator of *eLIBRARY.ru* is slightly higher than that of *Scopus* and only slightly lower than that of *WoS*. This may suggest that the technology of working with SI on the resources that are being com-

pared is roughly the same for foreign and Russian users.

(9) The small number of internal and outgoing links for *eLIBRARY.ru* may be due to the lack of transparency of this information to the software tools used by the authors of this article.

(10) According to the number of incoming links, *CyberLeninka* is ahead of *eLIBRARY.ru* by almost an order of magnitude.

(11) The VINITI website shows a modest number of incoming links, considering the role that this resource is intended to play in Russian science.

(12) Most of the WMIs of the LSL website are very high, but its AVD indicator is only 42 s.

(13) The relatively high relevance of the *MARS* resource can be due to regular site visits by Russian librarians for professional needs.

(14) The relatively high values of most WMIs for the *NEICON* archive are also noteworthy. However, the AVD of this resource is low, 58 s. Thus, based on the ratio of the WMIs, this resource is similar to the LSL.

CONCLUSIONS

(1) We showed that given the development of information and telecommunication technologies, scientific content-aggregator websites play a key role in ensuring the availability of scientific information to interested individuals and organizations.

(2) We proposed an original classification of content-aggregator websites.

(3) We revealed the shortcomings of the existing approaches to the assessment of indicators characterizing content-aggregator websites and to the development of integral (generalized) estimates for groups of these indicators.

(4) We proposed a methodology for remote assessment of specific webometric indicators (WMIs) of websites, which uses free software and public-domain Internet services.

(5) We estimated the WMIs of three categories of content-aggregator websites: foreign multithematic citation and scientometric systems recognized by the Higher Attestation Commission (HAC) of Russia; foreign research abstract and full-text DBs of scientific materials not recognized by the HAC; and the main Russian multithematic websites that aggregate scientific information.

(6) The comparison of the obtained WMIs allowed us to draw conclusions regarding the popularity of foreign and Russian websites, the relevance of the materials posted on them, and the features of user activities with scientific information resources on the content-aggregator websites.

REFERENCES

1. Brumshtein, Yu.M., An analysis of the impact of information and communication technology on the structure of scientific and technical information in Russia, *Sci. Tech. Inf. Process.*, 2016, vol. 43, no. 4, pp. 257–267.
2. Lopatina, N.V., The information infrastructure of society: Problems of research and management, *Sci. Tech. Inf. Process.*, 2016, vol. 43, no. 2, pp. 95–98.
3. Syuntuyurenko, O.V., The evolutionary factors, management, and efficiency of information support, *Autom. Doc. Math. Linguist.*, 2016, vol. 50, no. 3, pp. 117–125.
4. Il'yasov, B.G., Gerasimova, I.B., Karamzina, A.G., and Suyargulova, Yu.R., A model of accumulation of knowledge in a scientific school as a result of self-organization of the process of information exchange, *Tr. SPIIRAN*, 2016, no. 4 (47), pp. 144–162.
5. Zatsman, I.M., Internet electronic scientific document libraries: Structure, formal description, and nonverbal information retrieval, *Autom. Doc. Math. Linguist.*, 1998, vol. 32, no. 6, pp. 14–21.
6. Zelepukhina, V.A., The problem of reliability and objectivity of information within the scientific Internet community built on the principles of WEB 2.0, *Prikl. Vys. Tekhnol.*, 2013, no. 4, pp. 157–165.
7. Public resources: Government initiatives and practical solutions, *Univ. Kn.*, 2012, no. 4 pp. 20–27.
8. Gurov, A.N., Goncharova, Yu.G., and Bubyakin, G.B., Open access to scientific knowledge: Its state, problems, and prospects of development, *Sci. Tech. Inf. Process.*, 2016, vol. 43, no. 2, pp. 88–94.
9. Prokudin, D.E. and Chugunov, A.V., Network electronic repositories and aggregation and preservation systems in the concept of free access to scientific information, in *Integrirovannye tsifrovye resursy: Organizatsionno-tekhnologicheskie i nauchno-metodicheskie osnovy razvitiya. Sbornik nauchnykh trudov. Ser. "Elektronnaya biblioteka" Prezidentskaya biblioteka imeni B.N. El'tsina* (Integrated Digital Resources: Organizational-Technological and Scientific-Methodological Fundamentals of Development. Collection of Scientific Papers. Electronic Library Series, Yeltsin Presidential Library), Zhabko, E.D., Ed., St. Petersburg, 2015.
10. Björk, Bo-C., Open access to scientific publications—an analysis of the barriers to change?, *Inf. Res.*, 2004, vol. 9, no. 2, p. 170. <http://hdl.handle.net/10227/647>. Accessed July 4, 2017.
11. Laruk, O. and Garanovich, M.V., Assessment of access to scientific information for academic users on the Internet, *Global'nyi Nauchn. Potentsial*, 2014, no. 9 (42), pp. 135–138.
12. Prokhorova, A.M., The concept of site usability: Indicators and design standards, *Nov. Univ., Ser.: Ekon. Pravo*, 2016, no. 9 (67), pp. 87–90.
13. Laikova, A.A., Usability of the site: Principles and methods of assessment, *Aktual. Probl. Aviat. Kosmonavt.*, 2016, vol. 2, no. 12.
14. Egorov, V.S., Specifics of access to scholarly journal literature in the digital society: The history of the problem, *Nauchno-Tekh. Inf., Ser. I*, 2012, no. 11, pp. 18–23.

15. Arskii, Yu.M. and Bykov, V.A., The activities of VINITI RAS as a key organization of the commonwealth of independent states for the exchange of scientific and technological information, *Sci. Tech. Inf. Process.*, 2013, vol. 40, no. 2, pp. 101–108.
16. Kalinin, Yu.P., Khoroshilov, A.A., and Khoroshilov, A.A., Principles for the creation of a system for monitoring and analyzing the world flow of scientific and technical information, *Sist. Sredstva Inf.*, 2016, vol. 26, no. 1, pp. 139–165.
17. Mel'nikova, E.V., Enhanced functions of the modern STI system in Russia in the context of its improvement and of the development of the innovative orientation, *Sci. Tech. Inf. Process.*, 2014, vol. 41, no. 1, pp. 29–32.
18. Brumshtein, Yu.M., Vas'kovskii, E.Yu., and Kuanshkaliev, T.Kh., Searching for information on the Internet: An analysis of influencing factors and patterns of user behavior, *Izv. Volgogr. Gos. Tekh. Univ.*, 2017, no. 1 (196), pp. 50–55.
19. Biktimirov, M.R., Glebskii, V.L., Dolgov, B.V., and Polikarpov, S.A., Use of information technologies and infrastructures for the aggregation of scientific information. The experience of Canada, the Netherlands, and the United States, *Model. Anal. Inf. Sist.*, 2015, vol. 22, no. 1, pp. 114–126.
20. Lyagushkina, E.A. and Bogatov, V.V., Analysis of the effectiveness of fundamental scientific research: The publication aspect, *Naukovedenie*, 2017, vol. 9, no. 2. <http://naukovedenie.ru/PDF/10EVN217.pdf>.
21. Red'kina, N.S., Directions for the development of web analytics tools, *Nauchno-Tekh. Inf., Ser. 2*, 2017, no. 5, pp. 5–10.
22. Vas'kovskii, E.Yu. and Brumshtein, Yu.M., System analysis of the functionality of the counters of site traffic, *Priklasp. Zh.: Upr. Vys. Tekhnol.*, 2015, no. 3, pp. 45–58.
23. Guskov, A.E., Bykhovtsev, E.S., and Kosyakov, D.V., Alternative Webometrics: Study of the traffic of the websites of scientific organizations, *Sci. Tech. Inf. Process.*, 2015, vol. 42, no. 4, pp. 274–289.
24. Thelwall, M., Data Science Altmetrics, *J. Data Inf. Sci.*, 2016, vol. 1, no. 2, pp. 7–12. doi 10.20309/jdis.201610
25. Torres-Salinas, D., Cabezas-Clavijo, A., Jimenez-Contreras, E., Altmetrics: New indicators for scientific communication in Web 2.0. arXiv:1306.6595v1 [cs.DL]. doi 10.3916/C41-2013-0510.3916/C41-2013-05
26. Zahedi, Z., Costas, R., and Wouters, P., How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. arXiv:1404.1301v1 [cs.DL]. doi 10.1007/s11192-014-1264-010.1007/s11192-014-1264-0
27. Golitsina O.L., Kupriyanov V.M., Maksimov N.V. Information and technological solutions applied for knowledge-management tasks, *Sci. Tech. Inf. Process.*, 2015, vol. 42, no. 3, pp. 150–161.
28. Tret'yakova, O.V. and Kabakova, E.A., Possibilities of using webometric analysis in assessing the site of a scientific institute, *Vopr. Territ. Razvit.*, 2014, no. 2 (12), p. 5.
29. Shokin, Yu.I., Vesnin, A.Yu., Dobrynin, A.A., Klimenko, O.A., and Rychkova, E.V., Mathematical modeling of information processes in the web space, *Vychislitel'nye i informatsionnye tekhnologii v nauke, tekhnike i obrazovanii (CITech-2015)* (Computational and Informational Technologies in Science, Engineering and Education (CITech-2015)), 2015, pp. 300–315.
30. Kosyakov, D.V., Gus'kov, A.E., and Bykhovtsev, E.S., Academic institutes of Russia in the mirror of webometrics, *Vestn. Ross. Akad. Nauk*, 2016, vol. 86, no. 11, pp. 1015–1025.
31. Ivanova, V.I., Translation of scientific publications as a means of popularizing research results in the global world, *Russ. Yazyk Kul't. Zerkale Perevoda*, 2015, no. 1, pp. 207–215.
32. Semenyuk, E.P., Globalization of the information area and its significance for mankind, *Sci. Tech. Inf. Process.*, 2015, vol. 42, no. 1, pp. 1–12.
33. Skovorodina, I.S. and Efremov, P.V., Technologies of creation, aggregation and use of scientific and educational content, *Nauchn. Period.: Probl. Resheniya*, 2013, no. 6 (18), pp. 31–34.
34. Aksent'eva, M.S., Kirillova, O.V., and Moskaleva, O.V., To the problem of citing in the Web of Science and Scopus the papers from Russian journals that have translated versions, *Nauchn. Period.: Probl. Resheniya*, 2013, no. 4 (16), pp. 4–18.
35. Halevia, G., Moedb, H., and Bar-Ilan, J., Suitability of Google Scholar as a source of scientific information and a source of data for a scientific evaluation-Review of the literature, *J. Inf.*, 2017, vol. 11, no. 3, pp. 823–834. doi 10.1016/j.joi.2017.06.005
36. Antopolskii, A.B., On the feasibility of the Russian national webometric index, *Sci. Tech. Inf. Process.*, 2014, vol. 41, no. 1, pp. 33–37.
37. Pechnikov, A.A., Lugovaya, N.B., Chuiko, Yu.V., and Kosinets, I.E., Development of tools for webometric research of hyperlinks of scientific sites, *Vychisl. Tekhnol.*, 2009, vol. 14, no. 5, pp. 66–78.
38. Blekanov, I.S., Sergeev, S.L., and Martynenko, I.A., Construction of thematically oriented web crawlers using the generalized kernel, *Nauchno-Tekh. Vedomosti S.-Peterb. Gos. Polit. Univ., Inf., Telekomun., Upr.*, 2012, vol. 5, no. 157, pp. 9–15.
39. Brumshtein, Yu.M., Vas'kovskii, E.Yu., and Gorbacheva, A.N., Internet sites of executive authorities in the Caspian regions of Russia: Analysis of nomenclature, functionality, and demand for resources, *Priklasp. Zh.: Upr. Vys. Tekhnol.*, 2016, no. 4 (36), pp. 52–68.
40. Osipova, V.A., NEB: The history, structure, and the new stage of development, *Univ. Kn.*, 2013, no. 10, pp. 64–66.
41. Eremenko, G.O., Elibrary.ru: Course on improving the quality of content, *Univ. Kn.*, 2016, no. 3, pp. 62–68.
42. Yurkov A.V. Applied scientometrics: Elibrary.ru vs Google Academy, *Vestn. Ross. Univ. Druzhby Narodov, Ser.: Inf. Obrazov.*, 2015, no. 2, pp. 21–27.

Translated by A. Kobkova

EXHIBIT 16



Home › FAQ › SimilarWeb vs Google Analytics

SimilarWeb vs Google Analytics

Last Updated: May 23, 2019 08:51

SHARE:   

The purpose of this guide is to look at why there are sometimes discrepancies between the figures that SimilarWeb shows for a given URL and the figures that the owners of that URL see from direct measurement data tools such as Google Analytics.

This follows on from our review of [SimilarWeb's Data Methodology](#) which you can find [here](#).

Direct Measurement vs SimilarWeb Data

Why Does My Direct Measurement Data Show Something Different to SimilarWeb's Data?

There are three main reasons why the data you get from direct measurement tools such as Google Analytics may differ from the data shown in the SimilarWeb Platform.

- Different Methodologies
- The Human Touch
- Consistency

Different Methodologies

Most businesses use direct measurement tools, such as Google analytics, to analyze traffic to their own domains. While the technology is usually similar from one tool to the next, the data often varies. This is because of different methodologies used to calculate sessions, session times, and other standard metrics. For example, some methodologies deduplicate visits while SimilarWeb does not. Some tools discount or remove **bot** traffic, SimilarWeb does not.

The Human Touch

Trackers and analytics codes can be implemented on websites in different ways. Many tools rely on an imperfect human element. This can result in under-reporting pages where a code has not been



Consistency

When comparing numbers shown by SimilarWeb to numbers from a direct measurement tool, you need to make sure that your parameters are consistent.

- Are you comparing the same time frame?
- Are you comparing the same devices? E.g., desktop, mobile or **both**
- Are you comparing the same metrics? E.g., unique visitors vs visits

SimilarWeb uses a consistent methodology across all of our estimations, measuring all sites in a consistent unified manner.

Using SimilarWeb and Direct Measurement Data Together

Rather than competing against one another, direct measurement data complements SimilarWeb helping to calibrate for bias and move from sample to estimation by transforming data sources into intelligent estimations across all sites.

Next

What updates were made to the SimilarWeb extension?

Did you find this article helpful?

Yes, thank you!

Not really



RESOURCES

ABOUT US

OUR OFFICES

Follow us on:

+972-3-5447782

Derech Menachem Begin 121,
Tel Aviv-Yafo, 6701203, Israel

SIMILARWEB EXTENSION

[Categories](#) [Countries](#) [Privacy](#) [Terms](#) [Connect Google Analytics](#)

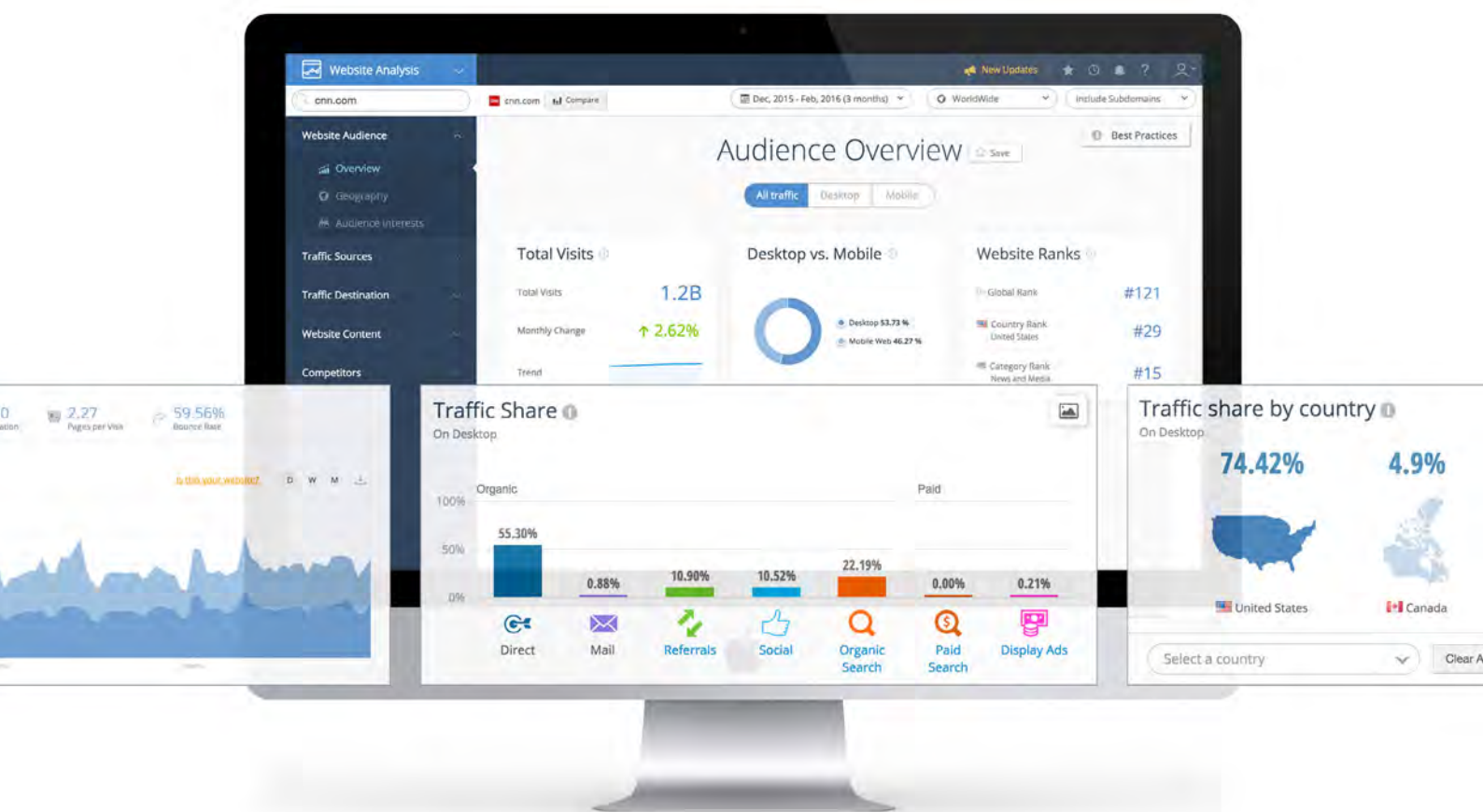
© SimilarWeb LTD 2019 All Rights Reserved

EXHIBIT 17



SimilarWeb vs. Direct Measurement

What is the Difference?



Understand the differences between SimilarWeb and direct measurement tools to gain the most value out of your digital market intelligence.



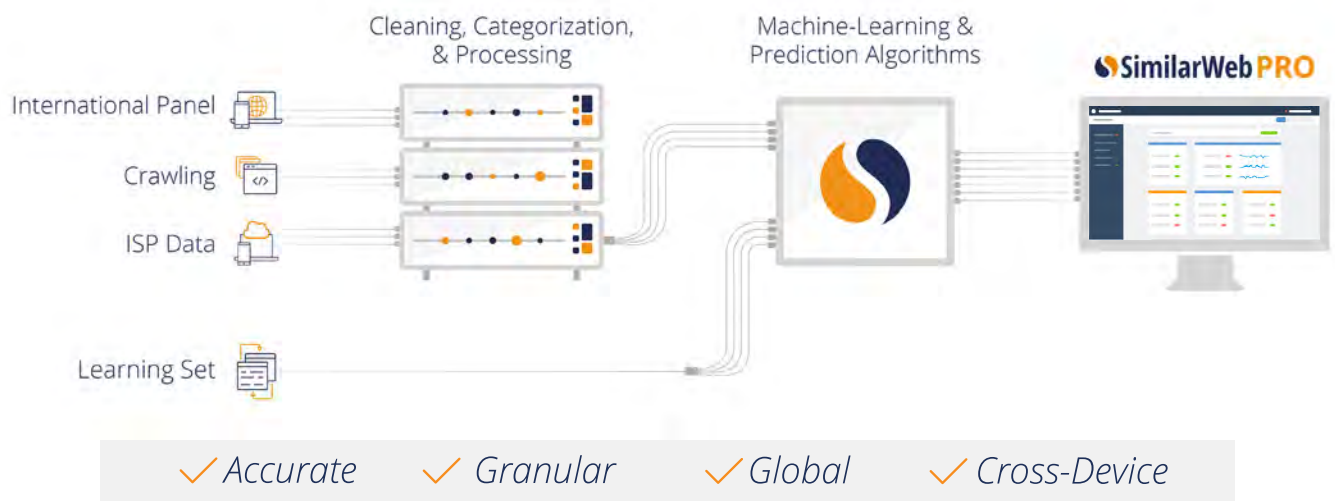
What is the difference between SimilarWeb market intelligence data and direct measurement tools?

- Direct measurement tools (e.g., Google Analytics) provide insight into your own site's metrics. SimilarWeb provides competitive market intelligence data for any industry, website or app - not just your own.

How does SimilarWeb's data collection differ from that of direct measurement tools?

- Direct measurement tools usually retrieve their data based on embedded codes on web pages or mobile apps.
- SimilarWeb collects anonymous clickstream data from a diverse panel of users, and employs unique algorithms to estimate overall metrics for web and apps. To improve the accuracy of our estimations, SimilarWeb calibrates its estimations using a learning set comprised of hundreds of thousands of directly measured websites and apps.

SimilarWeb Data & Methodology





- SimilarWeb's robust collection process enables capturing data that direct measurement tools miss, such as: data blocked by ad filtering, missed data due to Javascript issues, Google "not provided" keywords and secure (https:) to non-secure (http:) referral loss.
- Unlike other traffic analysis tools, SimilarWeb does not rely on cookies for counting unique visits, which is considered an unstable technology susceptible to being manually or automatically deleted.

Why might SimilarWeb data and direct measurement data differ from one another?

- When it comes to online measurement, there is normally up to a 20% discrepancy between analytics tools. Since no two measurement tools will produce exactly the same data, it is quite normal to see similar disparities of absolute numbers for SimilarWeb data.
- Google Analytics, Adobe Analytics, and other site-side analytics platforms are bespoke solutions which allow for custom setup (e.g., how an interaction is defined, how long is a session time, bot or spider filterings, domain/country combinations, etc.). These variations can result in differences between the absolute numbers or traffic source distributions that SimilarWeb shows compared to your site-side analytics.
- With SimilarWeb, you can trust that the **same measurement methodology is applied to all sites within the PRO in order to ensure a consistent baseline.**
- For this reason, SimilarWeb's traffic sources are classified according to the same uniform methodology across the platform.



SimilarWeb Traffic Sources Classification



Direct

- Address Bar
- Bookmark
- Autofill
- External Applications/ Software (Wordpad, Outlook, iMail)
- Pop Up Ads



Search

- Search Results (Google, Bing, Yahoo, etc.)
- Search Partners
- Search Channels
- Search Ads (Under Paid)



Mail

- Webmail (Browser-based)



Display Ads

- Ad Networks
- Recommendation Engines
- Taboola / Outbrain
- Tracking Domains



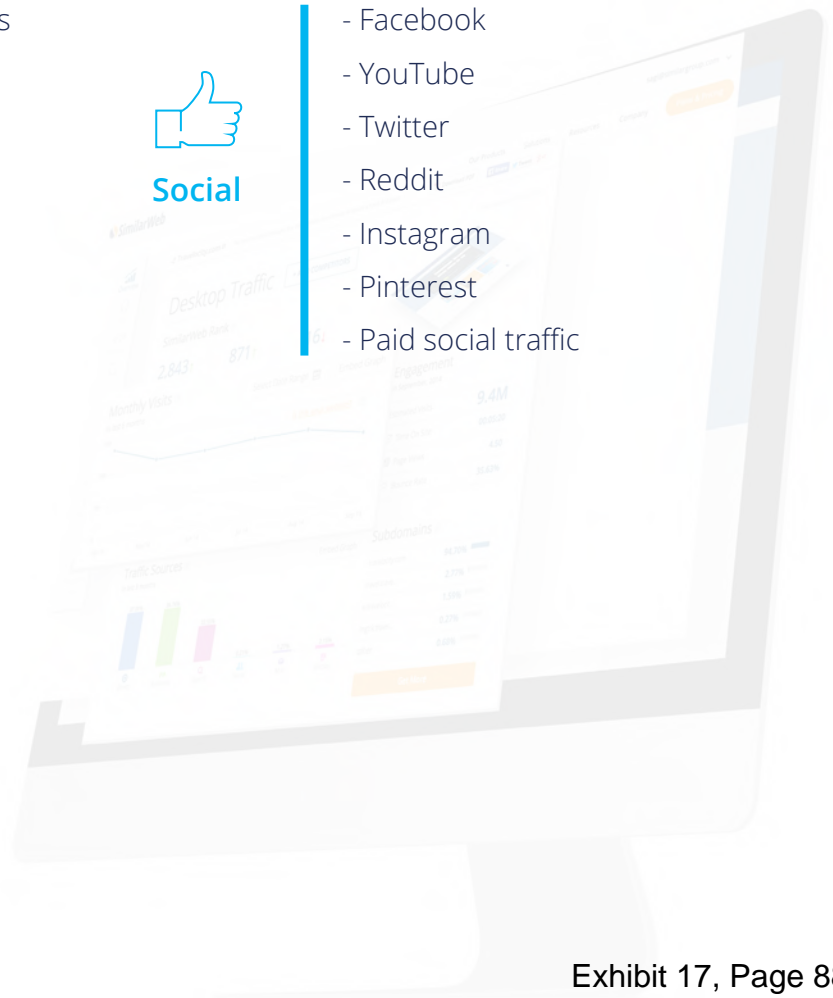
Referrals

- Links from other domains
- Affiliates
- Partnerships
- Content Marketing
- News Coverage
- Yelp
- Direct Media



Social

- Facebook
- YouTube
- Twitter
- Reddit
- Instagram
- Pinterest
- Paid social traffic





So, how can your internal analytics be used in conjunction with SimilarWeb's insights?

- Benchmark traffic share and performance - compare and monitor how each site stacks up for its share of industry traffic, engagement metrics, and channel-specific performance.
- Focus on the trends - analyse how visits, engagement metrics or traffic source distribution change month-to-month or year-over-year. Use this to spot changes in performance trends and shifts in marketing strategies.
- Look for the white spaces - SimilarWeb brings to light strategic insights and digital growth opportunities for any site in the world that you cannot see with your own analytics.

About SimilarWeb

SimilarWeb is a global cross-device market intelligence company used by 100,000s of businesses worldwide, including Google, eBay, L'Oréal and United Airlines, to discover, decide and deploy their digital strategy. SimilarWeb provides insights for every website, industry, and mobile app worldwide.

To learn more about SimilarWeb's data and methodology, [click here](#)

EXHIBIT 18

| Links | Title | Date | Cleaned_Links | Article_Classification | Date_Classification |
|---|---|-----------|---|------------------------|---------------------|
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | A tale of two Elons: Why this superhero should get banned from Twitter (but wor't) | 7/6/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk has revealed his Baseline Pedophile Claims | 7/7/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | Taiwan |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk Calls 'Key Player in Thai Cave Rescue a 'Pedo' | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Sorry pedo guy: Elon Musk lashes out at British cave rescue diver who criticized him | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | Canada |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Sorry pedo guy: Elon Musk lashes out at British cave rescuer Tesla Investor, Then Deletes 'tweet | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | Pakistan |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Sorry pedo guy: Elon Musk lashes out at British cave rescue diver who criticized him | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | Canada |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk calls British diver in Thai cave rescue a 'pedo' in Baseline attack | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk Calls British diver in Thai Soccer team rescue a Pedophile | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Sorry pedo guy: Elon Musk lashes out at British cave rescue diver who criticized him | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | Canada |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk Defends His Rejected Mini-Sub Plan for Thai Cave | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk calls British diver who helped free Thai schoolboys 'pedo guy' in Twitter clash | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | United Kingdom |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk calls British diver who advised Thai cave kid rescue 'pedo' on Twitter - No one wanted his little | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk calls British diver who helped free Thai schoolboys 'pedo' on Twitter - No one wanted his little | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk lashes out At Thai Cave Rescuer, Baselessly Calls Him A 'Pedo Guy' | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk lashes out At Thai Cave Rescuer, Baselessly Calls Him A 'Pedo Guy' | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk lashes out At Thai Cave Rescuer, Baselessly Calls Him A 'Pedo Guy' | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | 0 Pedito | |
| https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | Elon Musk lashes out At Thai Cave Rescuer, Baselessly Calls Him A 'Pedo Guy' | 7/15/2018 | https://thencr.com/2018/07/06/a-tale-of-two-elon-why-this-superhero-should-be-banned-from-twitter-but-wor/ | | |

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

Exhibit 18, Page 97

[illegible]

EXHIBIT 19

In Two Parts—24 Pages.
PART I—TODAY'S NEWS—IN THIS

The Los Angeles Times

Liberty Under Law—Equal Rights—True Industrial Freedom

SATURDAY MORNING, FEBRUARY 6, 1915. PRICE 5 CENTS

HERO FUND DEFENDED.

Carnegie a Witness in New York.

And John D. Rockefeller, Sr., Comes to Unpopularity in His Testimony.

World's Robert Mrs. Tell Why They are Getting Along Well.

Each Told His is Trying to Do Something Good for All Mankind.

SO IT PAYS TO BE UNWY.

Marshall P. Wilder, the Author, Lamented His Fate.

THE NEW YORK TIMES, Feb. 5.—(Special Telegrams.)—The New York Times, in its editorial of Feb. 4, has taken issue with the Los Angeles Times, which has published a series of articles on the hero fund, and the fact that the fund is being used for the benefit of the hero fund.

FLOOD PERIL INCREASES.

Blizzard Raging in Many States.

Cincinnati Union Decried in Its Fight to Stop Floods.

Train Wrecked in Rhode Island, Locomotive Blasted in Collision.

Police Post of Snow in Town, Earlier Part.

SEE A DANGER TO UNITED STATES IN BLOCKADE OF BRITISH ISLES.

Likelihood of a Protest to German Government Reported Under Discussion in Washington.

New Situation in International Relations Causes Some Worry and Wonder of the White House. State Department, by Joint Official Visit of German Embassy, Statement Before Taking Action.

THE WORLD'S NEWS

IN TODAY'S TIMES.

THE EVERETT REVIEW OF FEBRUARY 6. (1) Floods in Cincinnati. (2) German Embassy's Declaration of War as a Measure to American Shipping. (3) John D. Rockefeller, Sr., and Andrew Carnegie in the Witness Stand in New York. (4) Foreign Legation May Leave Mexico City. (5) Court Arrives at the Front: Russian House Offensive. (6) Congress. (7) Blizzard in the Middle West.

RUSSIANS RESUME OFFENSIVE DIRECTLY WEST OF WARSAW.

BY ATLANTIC CABLE AND R. E. J.

WARSAW, Feb. 5. (R. E. J.)—The Russian offensive against the German army, which was resumed on Feb. 4, has been reported by the Russian press to have been successful. The Russian army is reported to have advanced directly west of Warsaw, and to have captured a number of German positions. The Russian press also reports that the German army is in a state of retreat, and that the Russian army is in a position to capture Warsaw.

FOREIGN LEGATIONS READY TO WITHDRAW FROM MEXICO.

BY ATLANTIC CABLE AND R. E. J.

MEXICO CITY, Feb. 5. (R. E. J.)—The foreign legations in Mexico City are reported to be ready to withdraw from the city, in the event of a declaration of war between the United States and Germany. The legations are reported to be in a state of readiness, and to be prepared to leave the city at any time.

EXHIBIT 20

8/11/2019

Website Performance



foxnews.com



breaking news, latest news and current news from foxnews.com. breaking news and video. latest current news: u.s., world, entertainment, health, business, technology, politics, sports. Related Mobile App:

Global Rank #122

Country Rank #25
United States

Category Rank #21
News and Media

FOX NEWS more
Website Audience



Total Visits ⓘ
May 2019 - Jul 2019 ⓘ World

1.120B
3.80% from last month

Device Distribution ⓘ
May 2019 - Jul 2019 ⓘ Wor



Monthly Visits 373.3M
Monthly Unique Visitors
Avg. Visit Duration 00:06:20
Pages / Visit 3.27
Bounce Rate 49.86%

https://pro.similarweb.com/#!/website/worldwide-overview/foxnews.com/*999/3m?webSource=Total



1/5

8/11/2019

Website Performance

Geography

8/11/2019

Website Performance

Marketing Channels

Referrals

Search

8/11/2019

Website Performance

Social

64.15%

12.10%

12.00%

8.4

8/11/2019

Website Performance

Display Advertising

Outgoing Links

Outgoing Ads
